

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior,
and Problems of Aging

EDITED BY
CARL MURCHISON

JOHN E. ANDERSON
University of Minnesota

NANCY BAYLEY
Nat. Institutes of Health

JAMES E. BIRREN
Nat. Institutes of Health

WILLIAM E. BLATZ
University of Toronto

S. K. BOSE
University of Calcutta

CHARLOTTE BÜHLER
Univ. Southern California

CYRIL BURT
University of London

LEONARD CARMICHAEL
Smithsonian Institution

WAYNE DENNIS
Brooklyn College

SANTE DE SANCTIS
R. Università di Roma

WILMA T. DONAHUE
University of Michigan

ARNOLD GESELL
Yale University

ROBERT J. HAVIGHURST
University of Chicago

WILLIAM HEALY
Judge Baker Foundation

JOHN E. HORROCKS
Ohio State University

J. McV. HUNT
University of Illinois

WENDELL E. JEFFREY
University of California

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

ARTHUR T. JERSILD
Columbia University

BUFORD JOHNSON
Johns Hopkins University

HAROLD E. JONES
University of California

TRUMAN L. KELLEY
Harvard University

WILLIAM KESSEN
Yale University

ROBERT W. KLEEMEIER
Washington University

HEINRICH KLÜVER
University of Chicago

NORMAN L. MUNN
Bowdoin College

BERNICE L. NEUGARTEN
University of Chicago

TORAO OBONAI
Tokyo University

YOSHIKAZU OHWAKI
Tohoku University, Japan

HENRI PIÉRON
Université de Paris

ROBERT R. SEARS
Stanford University

NATHAN W. SHOCK
Nat. Institutes of Health

HAROLD W. STEVENSON
University of Minnesota

GEORGE D. STODDARD
University of Illinois

GEORGE G. THOMPSON
Ohio State University

C. J. WARDEN
Columbia University

VOLUME 95

1959

Copyright, 1959, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

Int. & Psych. Resear.

9.3.81

AUTHOR INDEX

Bayley, Nancy	83	Minnich, Robert E.	283
Bell, Richard Q.	83	Misiak, Henryk	13
Bloom, Leonard	3	Moyer, K. E.	255
Botwinick, Jack	137	Overall, J. E.	321
Brauchle, Robert P.	239	Pastore, Nicholas	157
Brinley, Joseph F.	137	Pfaffenberger, C. J.	145
Broadhurst, P. L.	217	Phaup, Minnie Rob	27
Brown, W. Lynn	105, 245, 321	Plutchik, Robert	171
Bunnell, B. N.	255	Pouliot, Samuel	13
Caldwell, Willard E.	27	Presley, W. J.	251
Carr, Richard M.	245, 321	Pritchard, Sally Ann	41
Cassel, Russell N.	239	Riopelle, A. J.	251
Charles, Don C.	41	Robbin, Joseph S.	137
Clifford, Edward	45	Rosenberg, B. G.	211
Cofer, Charles N.	293	Ross, Sherman	261
Cowley, J. J.	187	Schaefer, Earl S.	83
Fisher, Alan E.	261	Schwesinger, Gladys C.	329, 333
Goldfarb, William	203	Scott, J. P.	145, 261
Gonzalez, Richard C.	293	Singh, S. D.	217
Gordon, Jesse E.	177	Sinha, M. M.	19
Griesel, R. D.	187	Sinha, S. N.	217
Harder, Donald A.	207	Stewart, Lawrence H.	111
Kronovet, Esther	171	Sutton-Smith, B.	211
Lewis, Donald J.	207	Warren, J. M.	19
London, Perry	283	White, R. K.	321
McDowell, A. A.	105	Wurtz, Kenneth R.	227

TABLE OF CONTENTS

A reappraisal of Piaget's theory of moral judgement	3
BY LEONARD BLOOM	
The measurement of negative after-images in first-grade boys and girls	13
BY SAMUEL POULIOT AND HENRYK MISIAK	
Interactions between learning sets in monkeys	19
BY J. M. WARREN AND M. M. SINHA	
Some aspects of the behavior of the ant in the maze under various conditions of stimulation	27
BY WILLARD E. CALDWELL AND MINNIE ROB PHAUP	
Differential development of intelligence in the college years	41
BY DON C. CHARLES AND SALLY ANN PRITCHARD	
Discipline in the home: A controlled observational study of parental practices	45
BY EDWARD CLIFFORD	
Development of a maternal behavior research instrument	83
BY EARL S. SCHAEFER, RICHARD Q. BELL, AND NANCY BAYLEY	
A comparison of normal and irradiated monkeys on an oddity-reversal problem	105
BY A. A. McDOWELL AND W. LYNN BROWN	
Relationship of socioeconomic status to children's occupational attitudes and interests	111
BY LAWRENCE H. STEWART	
Modulation of speed of response with age	137
BY JACK BOTWINICK, JOSEPH F. BRINLEY, AND JOSEPH S. ROBBIN	
The relationship between delayed socialization and trainability in guide dogs	145
BY C. J. PFAFFENBERGER AND J. P. SCOTT	
Perceptual functioning in the duckling	157
BY NICHOLAS PASTORE	
Studies of parent-child relations: A critique	171
BY ROBERT PLUTCHIK AND ESTHER KRONOVET	
SHORT ARTICLES AND NOTES	177
Sex differences in the independence training attitude— <i>n</i> achievement relationship (Jesse E. Gordon)	177
BOOKS	181
BOOKS RECENTLY PUBLISHED	181

Some effects of a low protein diet on a first filial generation of white rats	187
BY J. J. COWLEY AND R. D. GRIESEL	
A method for appraising children in in-patient psychiatric treatment: Ranking for normality	203
BY WILLIAM GOLDFARB	
Partial reinforcement following a long continuous reinforcement acquisition series	207
BY DONALD J. LEWIS AND DONALD A. HARDER	
A scale to identify impulsive behavior in children	211
BY B. SUTTON-SMITH AND B. G. ROSENBERG	
The effect of stimulant and depressant drugs on a measure of emotional reactivity in the rat	217
BY P. L. BROADHURST, S. N. SINHA, AND S. D. SINGH	
The expression of guilt in fantasy and reality	227
BY KENNETH R. WURTZ	
An assessment of the fakability of scores on the group Personality Projective Test	239
BY RUSSELL N. CASSEL AND ROBERT P. BRAUCHLE	
Manipulation of visually homogeneous stimulus objects	245
BY RICHARD M. CARR AND W. LYNN BROWN	
Observational learning of an avoidance response	251
BY W. J. PRESLEY AND A. J. RIOPELLE	
Effect of exogenous adrenalin on the startle response in the rat	255
BY K. E. MOYER AND B. N. BUNNELL	
The effects of early enforced weaning on sucking behavior of puppies	261
BY J. P. SCOTT, SHERMAN ROSS, AND ALAN E. FISHER	
Effects of differential instructions and anxiety level on discrimination learning	283
BY ROBERT E. MINNICH AND PERRY LONDON	
Exploratory studies of verbal context by means of clustering in free recall	293
BY RICHARD C. GONZALEZ AND CHARLES N. COFER	
The effects of food deprivation and restricted activity upon exploratory behavior of the rat	321
BY R. M. CARR, J. E. OVERALL, R. K. WHITE, AND W. L. BROWN	
BOOKS	
Sargant, W. <i>Battle for the Mind: A Physiology of Conversion and Brain-Washing</i> . (Reviewed by Gladys C. Schwesinger)	329
Anastasi, A. <i>Differential Psychology</i> , 3rd Ed. (Reviewed by Gladys C. Schwesinger)	329
	333
BOOKS RECENTLY PUBLISHED	
	337

A REAPPRAISAL OF PIAGET'S THEORY OF MORAL JUDGEMENT*

Department of Psychology, University of Natal, South Africa

LEONARD BLOOM

The growth of modern child psychology has been mainly the result of two influences: the practical demands of the educationalists, and the theoretical stimulus of the theory of evolution. The former brought about the technological pressure that led to the development of intelligence and aptitude tests, and the latter stimulated the study of the growth of the child as a biological problem. Psychologists distinguished the stages of childhood in the various functions of personality, as biologists and zoologists traced stages of development in their studies.

Of particular interest to the student of child and social psychology is the nature of the child's moral development, and one of the pioneers in the scientific study of these motives of children was Jean Piaget. He had no place for *a priori* moralising about the child's notions of right and wrong, and without allegiance to any school of psychology, he used empirical methods that permitted the repetition and testing of his enquiries, and the verification, rejection, or modification of his hypotheses.

Piaget's classic research, *The Moral Judgement of the Child* (12), appeared in England in 1932, and is based upon a simple situation, natural and familiar to the child subjects of the investigations: a game of marbles played in working-class districts of Geneva and Neuchatel. This close regard for what activity the child finds natural is an integral part of the early method of Piaget. He claims to have discerned a tendency in children to think of laws and rules in the most factual and literal way. The child is not concerned that he is only playing a game, but follows the letter of the law rigidly. There is an unquestioning acceptance of the logic and justice of a moral rule, even though the reflecting adult might think the rule is harsh, unjust, inappropriate, arbitrary, or foolish. The rule is part of the given world and can no more be changed than the tides can be meddled with or the rhythm of the seasons altered. This "moral realism" is more typical of the younger child; as the child grows older he becomes more flexible and more imaginative.

* Received in the Editorial Office on March 26, 1958.

Piaget next discusses "co-operation and the development of the idea of justice." He asserts that the child's notions of fairness and justice are independent of adult precept and influence, requiring "nothing more for (their) development than the mutual respect and solidarity which holds among children themselves." Gradually, a predominantly retributive and equalitarian justice is replaced by more rational and objective considerations.

Piaget claims that children pass through four successive stages in their play. In the *motor and individual* stage, the child plays individually and as his feelings and motor habits dictate. He tosses blocks about the room, he rattles pebbles in a tin, and collective rules are irrelevant to him. It is indeed difficult to distinguish in this behavior a game with rules in contrast with inchoate exploration of the physical world.

Next the *egocentric* stage, between two to five years onwards, in which the child first receives from outside himself the example of definite codes of rules. There is now a pattern to his playing. The child imitates the example of the rules that he notices, but he still continues largely to play by himself not troubling too much to find other children with whom to play, and without trying to establish spontaneously his own conventions and rules. He plays egocentrically, in a parallel fashion; all can win the games, no one bothers to define the rules, but they are followed as far as they are recognized.

Between the ages of seven and eight years the stage of *incipient co-operation* appears. The aims of competing and winning become important, and children are much concerned with clear definitions of the rule and protocol. But even now they are not entirely clear about the rules, and if they are asked individually to define rules they may give contradictory accounts.

At about 11 to 12 the child becomes concerned with the *codification of rules* and the details of play are well known by the whole group and carefully observed by its members.

These findings suggest the hypothesis that to very young children "a rule is sacred because it is traditional; for the older ones it depends on mutual agreement." Piaget continues to argue that for children from five to eight "social life and individual life are one," and the gradual diminution of conformity that takes place as the child grows is a correlate of increasingly clear and definite structuring of his personality. In effect agreeing with Kurt Lewin, he writes that "childish egocentrism is in its essence an inability to differentiate between ego and the social environment." However, as the child becomes less subject to the supervision of elders and adults, as he takes

part in more social groups and social situations, and as he becomes physically and emotionally like older children and adults, we see his awe diminishing and his individuality asserting itself.

Extending the child's attitudes to the rules of the game to his attitude to rules in general, Piaget shows that the responsibility for the breach of a rule is not judged after taking motives into account, but according to the results of the behavior. A child judges as naughtier another who accidentally breaks three cups, than one who deliberately smashes one. It would be interesting to refer this back to games and to ask, for example, if children regard as more serious the fault of a child who "loses" the game by two goals by his justifiable lack of skill, or one who loses it by one goal as a result of his deliberate inattention. Older children are less prone to judge naughtiness by the amount of "damage" done, and are correspondingly more ready to consider the offender's motives. The young child absorbs and adopts the adult way of looking at lies and damage, and because adult responses are often harsh and literal the child adopts harsh and literal notions of responsibility. Adults tend to present to the child cut-and-dried obligations before he is sufficiently mature emotionally and socially to assimilate them intelligently, i.e., to apply them to relevant situations; the child fails to understand the principles (if any) behind the obligations and the offences and cannot generalize to analogous situations. Instead the obligations become "ritual necessities," "taboos," and an "ineffectual deduction from the words spoken by adults." This is re-enforced by memory and experience: thought lags behind action, and once the child has experienced a dramatic situation of adult constraint he will continue to judge according to that earlier situation though it is no longer appropriate.

Similar considerations apply to lies. Not until Piaget questioned children of about 10 to 11 did he find evidence that children believed that to tell a lie you must tell an intentional falsehood. This shows "how completely external to the child's own mind the interdiction of lying still remains in the early stages." Thus: "when a material action is too closely attached to the lie the child has a certain difficulty in dissociating the lie so far as it is psychological action from the actual results of the concomitant external act."

In sharp contrast with his views of the origins of moral realism, Piaget holds the child's ideas of justice to be largely independent of the example of the adult. This contrasts too with the views of psychoanalysis on the moral development of the child. He distinguishes three periods in this development: (a) Lasting to seven or eight in which justice is subordinated to adult authority; (b) between about eight and 11, which is the stage of progressive equali-

tarianism; (c) from about 11 or 12 during which equalitarian justice is tempered by considerations of equity. In the first stage every punishment is accepted "as perfectly legitimate, as necessary, and even as constituting the essence of morality; if lying were not punished one would be allowed to tell lies." In the second stage punishment becomes do-as-you-would-be-done-by, equality increases, good behavior begins to be sought for its own sake independently of reward or punishment. Piaget does not, however, analyse the "reward" of adult approval and general benignity. How does the child learn to distinguish good from bad behavior unless there are "rewards" and "punishments," in situations in which he can recognize the appropriateness of the one or the other? In the last stage the child takes into account the circumstances of each case individually, and no longer are the same law and the same punishment applicable to all children in the same way.

Piaget was strongly influenced by the sociologist Emile Durkheim, and he concludes with some socio-psychological observations on child morality. There are, he argues, two types: the inflexible and group enforced as distinct from the objective and autonomous, analogous to the fundamental types of social relationship posited by Durkheim in his *On the Division of Labour in Society*. Essentially: ethics and morality can only be explained by reference to social organization, and changes in morality can only be understood by referring to changes in social structure and socially sanctioned values. Morality presupposes rules and laws transcending the individual; out of individual action emerges a code of morality explicable *sui generis* without reference to the personal wishes, goals, hopes, and acts of individuals. Child morality stems either from adult or from child constraint, and in either case the morality that emerges is social. Coupled with the behavioral conformity of the child and adult to a moral code is a subjective feeling of constraint, which is the root of the sense of guilt or shame, remorse, right and wrong, conscience. These subjective feelings, too, are the result of social values and pressure and not the contrary.

Piaget, despite his acute awareness of the power and autonomy of social determinants of behavior, is sharply critical of the exclusively social explanations of child psychology. Piaget argues that because there are two networks of social relations for children there are two types of child morality. There is the morality that arises from the network of child relationships among themselves, and that which arises from the relationships of adults to children. The child not only lives by the code of rules that the adults enforce, but by that enforced by his fellows. Though both these codes are social they reflect two quite different types of authority: that due to uni-

lateral respect (e.g., the attitude of the younger to the older child) and that due to mutual respect, and restraint is different in these two types. As contemporary society is tending to adopt the rule of co-operation in place of the rule of naked constraint, so in education coercion is giving way to "the habit of internal discipline, of mutual respect and of 'self-government.'"

In evaluating Piaget's contribution, it is tempting to confine the criticism to his inadequate sampling, lack of controls, haphazard reporting, and presentation. It must be admitted that his work shows these technical defects, but these objections in themselves would not completely fault Piaget's original and highly stimulating work. More serious objections lie in his apparent oblivion to the part that cultural and class factors may play in determining the child's judgements. These ignored factors may affect both the content of the child's beliefs, and the psychological processes of their development. Liu (10) found that native-born American children raised in America showed much greater severity of moral judgement and much less subtlety of analysis than Chinese-born Americans raised in America, and he considered that these marked differences were the result of the great intimacy of the Chinese-American family and the Confucian philosophy which encourages consideration of moral questions. Harrower (5) found that English middle-class children showed much more temperate and flexible moral judgements than their working class peers, and that both groups made fewer retributive and immature judgements than Piaget's sample. Other investigators such as Ball (2), MacRae (11), Himmelweit (6), Havighurst and Taba report similarly that Piaget's findings need to be modified to allow for class and cultural differences.

A further complication is that Piaget nowhere suggests the psychological dynamics of the learning of moral judgements; why it is, for example, that at one time the child learns from his peers and at another from adults. Piaget's discussions of the learning are essentially in a rather old-fashioned behavioristic stimulus-response framework, with the stimulus being social pressure in one form or another. The subtlety of the interrelationships between children and between children and adults is nowhere suggested, and from the Piaget and Durkheim point of view it is difficult to understand how a child might ever develop moral judgements that are not accepted by or acceptable to his group. How can we explain, understand, and help the anti-social child, who resists or for other reasons does not follow the lines of conduct and thought that his society provides? Who is not "constrained" like his fellows. It may be that by adult standards his moral judgement is of a higher ethical quality than that of his peers, he may for example show

a gentleness and forgivingness that is not to be expected from his age group, or he may commit anti-social behavior. The child does not learn moral judgement in the somewhat mechanical fashion that Piaget implies, but will only learn spontaneously and fruitfully if his personal relationships with his peers and adults are emotionally satisfactory to him, for it has been amply demonstrated by now that the child's emotional development (of which his moral judgement is one major aspect) depends upon the child's having adult models from whom he introjects his moral attitudes. He may simulate a group-sanctioned moral code, but so may a slave simulate what his masters desire him to. Unless he has learned to trust and to expect justice and morality from the adult world in very young childhood, he will not learn to make the conscious and active moral judgements that are most worthy of men.

The content of moral judgement is largely learned from the norms that society has developed, but since Piaget wrote his book, we have begun to consider in more detail the differences between children and adults in their mode or manner of making moral judgements. We are concerned more with such personality differences as rigidity and flexibility in making judgements than in specific manifestations of moral judgement, and in this too we see a reason to consider that the Piaget-behavioristic approach is inadequate to answer problems that are significant to modern child and social psychology. The *quality* of the emotional responses of the child to matters of conscience (or more mundane problems such as the rules of a game of marbles, which are often of gravest import to a child), severity of super-ego or laxness, gentleness or stringency and the like are emotional variables that mediate between social demands and ordinances, and are no less a result of personal relationships than of social constraint—a constraint that some children are ignorant of, and others acutely sensitive to, while yet others grudgingly capitulate with feelings of resentment, and others joyfully accept.

Piaget at once overestimates and underestimates the cognitive element in child behavior. He overestimates it in his reliance upon the questionnaire technic to give valid and reliable information about the child's behavior in the vagaries of every-day situations. He underestimates it in the implicit and questionable assumption that there are fundamental differences in the modes of thought of adults as compared with children. We are nowhere told the intelligence range of the sample that Piaget questioned in terms of results on a standard intelligence test, of which plenty existed in 1932. It is arguable that *IQ* is very directly connected with the child's ability to make conceptual judgements and enter into subtle conversations with erudite psychologists, and if Piaget's sample was of children of lower intelligence it is

not unexpected that many of their judgements seem crude by adult standards, illogical and undifferentiated. There is a mixture of logic and irrationality in both children and adults, but it cannot be assumed *a priori* that logic or irrationality are more typical of one "stage" than another. We have to equate the problems and situations for the intellectual capacity and experience of the individual, which does not tacitly admit that there are different stages in the capacity for logical thought: the child thinks both logically and illogically, but about many different things from those that the adult considers. This is largely a function of the different size and content of the child's phenomenal world. Take an adult outside his common experience and he too is prone to think irrationally. In this book as in his others, Piaget seems to have an implicit opinion that the complexity of childthought is much less than other child psychologists would think to be the case. A consideration of the verbal behavior of the child suggests that the child is from a very early age capable of quite complicated and adequate thought processes and differentiation, using by the age of eight all parts of speech and a very large vocabulary which suggests that he is capable of employing the higher mental processes if in a rudimentary form. We cannot but feel, therefore, that some of Piaget's results are due to his questioning children of lower *IQ*, or that he failed to obtain from them answers as well differentiated as children of their age could reasonably give.

The problems that Piaget touched upon in this enquiry were singularly important to children, and though Piaget is a skillful, sensitive, and thorough interviewer, it may be that if the questions were presented differently, e.g., indirectly, the responses would have differed. It is possible that the results of later investigators which contradict those of Piaget are due to this factor of uncertainty. Piaget, too, ignores the child's use of gesture and expression to convey his meaning. A child may reply "Yes," sardonically, doubtfully, as a full agreement . . . ; we are rarely told of the circumstance surrounding the answers that would permit to decide for ourselves precisely what the child meant. It is remarkable that so few children gave qualifications in their answers; if children are as rigid as Piaget occasionally suggests it seems highly probable that their answers would often begin with a phrase like: "It all depends . . .".

Modern psychology is much concerned with "stages" in development, but we cannot legitimately assume that there are clear-cut stages through which the child moves (effortlessly?), without conflict and doubt, unerringly absorbing the crystal-clear moral and social precepts and codes of rules that are presented to him by his fellows or by the adults with whom he has a

significant relationship. Piaget's system can incorporate the data from other cultures and social classes; this is a straight-forward problem of expressing development in general terms that can be demonstrated by different contents. For Piaget, as for the Freudians, it is arguable that the processes of development are the same the world over, because the bio-psychological problems of maturation are the same for all people; it is the social and outward expression of basic problems of maturation that differ. This consideration only requires Piaget to adapt his theory to take account of local factors. But Piaget does not consider to what extent the child's moral development and judgement is affected by conflict in moral codes and between them. To what extent can the child recognize and comprehend such conflicts? To what degree is the child unable to develop as an adequate and healthy moral being because of the riven nature of the world he is living in, or because of an inability or lack of opportunity to form the emotional attachments which are basic to emotional-moral development, making it (and other social-personal development) meaningful and relevant?

Conflicts in the child's moral world may come from differences between his adult companions' and his child fellows' values, or from conflict among the adults from whom he is learning. Adult conflict and confusion may come from conflict in the social world reflected in the uncertainty of his adult mentors, but it may also come from the inadequacy of the adults to understand or to manipulate the world for their emotional satisfaction: they may be intellectually or emotionally unable to present to the child an accurate picture or model, because they are themselves unable to recognize it. This lack of attention to conflict in the child's world of moral (and other) judgements is perhaps derived from Piaget's tendency to emphasise the intellectual and logical aspects of thinking, which became more marked in his later work. This change in approach has been neatly summed by James Anthony. Piaget's "happy animistic child, chatting so unselfconsciously and charmingly on the mountains of the Salève or by the shores of Lake Geneva, has been transformed into a mathematical automaton, a cybernetic machine, with, as Piaget proudly proclaims, a negative feedback, the only one of its kind in developmental psychology." Piaget's early work paralleled in the field of cognition Freud's work in the field of the affective life. His early research was marked with a sensitivity which converted the verbal interview technic (despite its many drawbacks) into a subtle and useful instrument of investigation. It is possible that had Piaget continued to perfect his observations, he would have been led to appreciate the conflictual and affective factors in child life, and with his interests in sociology and

psychology, he would have contributed much to a more precise factual definition of the parts played by social and psychological, and affective and cognitive factors in child behavior. This would have done much to integrate three loosely connected disciplines: sociology, psychology, and psychoanalysis. But the opportunity was missed and Piaget continues to pursue a predominantly behavioristic path, unperturbed by the crying need to explore for the child the matrix of meaning and significance that has been tentatively explored for adults by such psychologists as Carl Rogers and Snygg and Combs.

However, Piaget's achievements are considerable, and his findings not to be ignored but used as a basis for further investigation, particularly from an affective point of view. Piaget showed that it was no more difficult to carry out a behavioristically oriented research with children than it was with adults, to discover psychological variables behind behavior. In so doing he stimulated and provoked a great deal of empirical research and observation. He converted child study from a branch of general psychology that was largely anecdotal and lacking in scientific rigor, isolated and ignored, to a central position comparable in importance to the once similarly neglected study of abnormal psychology. In this he has looked afresh at basic problems such as the nature of learning, perception, motivation, and intelligence, and from simple (but ingenious) experiments has drawn far-reaching theoretical conclusions, adding to the small stock of psychological theory based upon empirical research.

Piaget made a highly laudable pioneering attempt to describe in a more or less natural environment the behavior of normal children, avoiding the aridity of the strict Watson laboratory investigation, and the loose, anecdotal reports of the early child psychologists. In his early work Piaget was very closely and obviously interested in the child's spontaneous behavior, which makes his books a gold-mine for the prospector for unresolved problems in child psychology, and of value to the practitioner such as the infant teacher who is about to embark upon a considered study of child psychology. It is impossible to imagine that child psychology can ever return to the quasi-scientific methods of the period before Piaget made his contribution.

SUMMARY

The main argument of Piaget's *The Moral Judgement of the Child* is given in outline, in particular the gradual development of child morality from moral realism and objectivity, to a relativistic and subjective liberal

judgement. Piaget's argument that the child's moral judgement is a result of social pressure, from children or from adults, is shown to derive from Emile Durkheim. The argument of Piaget is criticised as being too exclusively based upon social causation, ignoring the rôle of conflict and affectional relationships in the child, and for a too rigid adherence to the theory of stages. Piaget is a pioneer in the observational and scientific approach to child study, for his fruitful stimulation of theory with empirical research, and for his raising child psychology to a status central within general psychology.

REFERENCES

1. ANTHONY, J. The system makers: Piaget and Freud. *Brit. J. Med. Psychol.*, 1957, **30**, Part 4, 1-12.
2. BALL, A. A test of group conformity in the moral judgments of children. *Brit. J. Educ. Psychol.*, 1955, **25**, 129-130.
3. BERLYNE, D. E. Recent developments in Piaget's work. *Brit. J. Educ. Psychol.*, 1957, **27**, 1-12.
4. BROWN, A. W., *et al.* Influence of affectional family relationships on character development. *J. Abn. & Soc. Psychol.*, 1947, **42**, 422-428.
5. HARROWER, M. R. Social status and the moral development of the child. *Brit. J. Educ. Psychol.*, 1934, **4**, 75-95.
6. HIMMELWEIT, H. T. A study of the attitudes, value systems, and behavior of young adolescents belonging to different social status groups. *Bull. Brit. Psychol. Soc.*, 1953, **20**, 7.
7. INHELDER, B. Developmental psychology. In *Annual Review of Psychology*. Palo Alto, Calif.: Annual Reviews, 1957.
8. ———. Criteria of the stages of mental development. In *Discussions on Child Development* (J. M. Tanner and B. Inhelder, Eds.). London: Tavistock, 1956.
9. ISAACS, N. The significance of Piaget. *Bull. Brit. Psychol. Soc.*, 1958, **34**, 67.
10. LIU, C. H. The influence of cultural background on the moral judgment of the child. Ph.D. Dissertation, Columbia University, 1950.
11. MACRAE, D. J. A test of Piaget's theories of moral development. *J. Abn. & Soc. Psychol.*, 1949, **49**, 14-18.
12. PIAGET, J. *The Moral Judgment of the Child*. London: Kegan Paul, 1932.

Department of Psychology
University of Natal
Durban, South Africa

THE MEASUREMENT OF NEGATIVE AFTER-IMAGES IN FIRST-GRADE BOYS AND GIRLS*

Department of Psychology, Fordham University

SAMUEL POULIOT AND HENRYK MISIAK

A. INTRODUCTION

While studies of eidetic imagery have been done on younger children, very little work on after-images has been done with children. One notable exception is the study reported in 1945 by Morsh and Abbott (2) who investigated the after-image behavior of more than 700 subjects. The blink method with the dark adapted eye was used on some subjects and with others the fixation of pictures, a technique developed by Klüver. Morsh and Abbott found no support for the existence of a special "eidetic" ability and no clear-cut difference between eidetikers and non-eidetikers. They report: "After-image abilities of a random group of subjects appear in a continuum." Although the subjects ranged in age from six to adulthood, the authors do not report how many younger children, for example, children aged 6 or 7, there were, nor do they report the results for the young children. They do state that "below the age of eight it is extremely difficult to get a reliable report of imagery."

B. PROBLEM

In studying after-images in children one of the principal difficulties is methodological. Therefore, one of the main objectives of this study was to develop a procedure suitable for young children and subsequently to employ it in the investigation of after-images as experienced by six-year-old children. More specifically, answers to the following questions were sought: (a) Do six-year-old children experience negative after-images and can they report accurately the color seen? (b) What is the variability in latency, duration, and number of reappearances of after-images within the group? (c) Can these variables be measured and how reliable are such measures? (d) Are there any sex differences regarding latency, duration, and the number of reappearances in after-images?

* Received in the Editorial Office on April 1, 1958.

C. SUBJECTS

Forty American-born white first-grade pupils between the ages of six years, four months, and seven years participated in this experiment. There were two groups, 20 boys and 20 girls, each group having a mean age of six years, nine months. All subjects were examined for freedom from visual defects including color-blindness.

D. PROCEDURE

The following three preliminary tests were employed to check on color-blindness, color identification, and identification of pictures respectively:

1. *The Ishihara Test of Color-Blindness.*

2. A color naming test to check on the child's ability to identify primary and complementary colors. It consisted of 10 sheets of paper, the colors of which were white, black, red, yellow, green, blue, orange, purple, gray, and magenta.

3. A test of identification of familiar animals and birds. Designed specifically for this study, the test was made up of pictures obtained from a series of children's books published by the Merrill Company. These achromatic representations of a cow, squirrel, cat, pig, horse, duck, lamb, rooster, chicken, turkey, and pigeon were pasted on white paper.

For the specific purpose of eliciting after-images three stimuli were employed: a magenta rabbit, a blue-green squirrel, and a yellow chicken. These colors were selected because it was desired to elicit after-images of primary colors, green, red, and blue, which would be easy for the young subjects to describe. All drawings were mounted in the center of a sheet of gray paper (8" x 11"), a fixation point being indicated on each stimulus. Nine similar stimuli were used in the foreperiod as practice trials: a blue lamb, magenta duck, green hen, blue kitten, blue-green dog, yellow goose, blue-green puppy, and a green pig. A stiff gray cardboard (9" x 12"), with a focal point in the center served as a cover for the stimulus after the exposure time had elapsed, and at the same time as an after-field for the after-image.

For the presentation of these stimuli the child was seated at a school desk (to which the seat was attached) with a flat top and located near the wall. A chin rest was clamped to the desk to prevent the child from moving his head. Directly in front of him, 15 inches away, a rack (15" x 24") attached to the desk held the stimulus and the gray cardboard cover at the end of the exposure time. The wall behind the rack was covered with gray paper. Between the chin rest and the rack a snowy 40-watt light illuminated the

stimulus and the after-field. A black shade protected the subject's eyes from direct light.

The instruments were: a Gray Laboratory Universal Timer, Type 165 (measuring time in seconds); two Standard Electric Time Company chronoscopes, Type S-1 (measuring time in hundredths of a second), operated on 115 volts *AC* with a 6-volt *DC* clutch, the Heathkit Battery Eliminator delivering 6 volts *DC* for the clutch of the chronoscopes, and a relay.

Each subject was tested individually in two sessions nine days apart. During the first session the subject took the three preliminary tests for color-blindness, color identification, and object naming, and was instructed and trained how to behave in the experiment. He was taught how to look steadily at the pictures, fixate the focal point, avoid blinking, and not move. When the stimulus was covered by the neutral after-field, the subject had to fixate on the new focal point. The most important was to train the child to report the appearance of the image and its disappearance, and continue to look for the reappearance of the image. Before the actual experiment each child had to complete two satisfactory trials.

The experiment proper consisted of three trials. Before the first, all lights were turned out. The experimenter placed the magenta rabbit on the rack. At the signal "Ready" he pressed the switch which started the Gray timer set for a 30-second exposure of the stimulus. During the exposure period the experimenter watched the child's eyes and kept instructing the child to continue to look at the picture.

At the end of the 30 seconds the experimenter swiftly slid the neutral gray background over the stimulus card while the timer automatically started the first chronoscope which measured the latency period. The chronoscope continued to register until the child reported an after-image, at which point the experimenter pressed the stopkey and kept the lever down. This automatically stopped the first chronoscope and started the second which measured duration. When the subject reported disappearance of the image, the experimenter released the lever. When the after-image reappeared, the experimenter depressed the stopkey and marked a tally on the score sheet.

When the first trial was completed the experimenter recorded the latency time as registered on the first chronoscope and the duration time as indicated on the other. The child was then asked the color, size, and shape of the image he had seen. Thereafter the second trial, a blue-green squirrel, was given and lastly, the third trial, a yellow chicken. Nine days later, the same procedure was employed for the second experimental session. The same stimuli were used in the same sequence.

E. RESULTS

The data consisted of the latency period and the duration of the after-images measured in seconds, and of the number of reappearances of after-images. These data were collected in two sessions. Analysis indicated that the measurements were reliable and consistent. This fact justified the averaging of the scores obtained in the first and second sessions. Table 1 presents

TABLE 1
SEX DIFFERENCES IN LATENCY AND DURATION IN SECONDS, AND NUMBER OF
REAPPEARANCES OF AFTER-IMAGES (N = 40)

	Latency		Duration		Number of reappearances	
	Boys	Girls	Boys	Girls	Boys	Girls
Mean	2.51	2.23	19.25	16.70	4.14	3.26
SD	1.30	0.56	21.52	8.98	4.72	3.28
$M_B - M_G$	0.28		2.55		0.98	
t	0.84		0.54		0.62	

the mean scores in latency and duration in seconds and the number of reappearances of after-images for both sexes.

Inspection of Table 1 shows that the girls' scores tend to be lower than the boys' in latency, duration, and reappearances, and that their variability is smaller. The difference between the boys' and girls' scores was not significant. Similarly, Morsh and Abbott did not find any significant sex differences in their eight-year-old children.

The effect of color on the latency, duration, and reappearances was also studied. The results are presented in Table 2. When the duration of after-

TABLE 2
AVERAGE LATENCY AND DURATION IN SECONDS AND NUMBER OF REAPPEARANCES OF
AFTER-IMAGES FOR THE THREE COLORED STIMULI (N = 40)

Stimulus	Latency		Duration		Reappearances	
	Boys	Girls	Boys	Girls	Boys	Girls
Magenta	2.31	2.12	17.32	21.82	3.60	3.90
Blue-Green	2.64	2.19	20.97	14.90	4.40	3.10
Yellow	2.60	2.43	19.51	12.64	4.30	2.80

images is considered, it is found that the after-image of the yellow stimulus, blue in color, had the shortest duration for both sexes. The two other colors showed sex differences. The magenta stimulus produced the longest duration of after-image, green in color, in girls whereas blue-green gave the longest after-image, red in color, in boys. No explanation can be offered for these

differences. Color preference could not be invoked here in view of the study of Katz and Breed (1) who did not find any significant sex differences in children in the first grade.

F. SUMMARY

The procedure for measuring latency, duration, and reappearances of after-images of young children was developed. It was tested on 40 American-born white six-year-old school children of both sexes and was found successful and capable of giving reliable results.

In the experiment no significant differences were found between boys and girls with respect to latency, duration, and number of reappearances. However, the girls had consistently lower scores and less group variability than the boys. The green after-image was of longest duration for the girls, whereas for the boys the red after-image had the longest duration.

REFERENCES

1. KATZ, S. E., & BREED, F. S. The color preferences of children. *J. Appl. Psychol.*, 1922, **6**, 255-266.
2. MORSH, J. E., & ABBOTT, H. D. An investigation of after-images. *J. Comp. Psychol.*, 1945, **38**, 47-63.

Department of Psychology
Fordham University
New York 58, New York

INTERACTIONS BETWEEN LEARNING SETS IN MONKEYS*¹

Departments of Psychology, Stanford University and Gorakhpur University

J. M. WARREN AND M. M. SINHA

A. PROBLEM

The experiments described in this paper were designed (a) to investigate the effects of reinforcing single stimuli upon subsequent pattern discrimination by monkeys, and (b) to study the solution of the double alternation problem by a large group of experimentally sophisticated rhesus monkeys. These objectives were incompletely realized, but the partial failure of these researches suggested several fundamental questions regarding the interactions between learning sets.

The questions which the present researches were designed to answer may be stated briefly. Lashley and Wade (7) found that 200 reinforcements of a single stimulus did not produce any positive transfer effects in later discrimination learning involving paired presentations of the rewarded stimulus and an unfamiliar stimulus. Other workers (2, 6, 17), however, have failed to duplicate these results. The discrepant results can be attributed to either of two procedural differences. Lashley and Wade studied experimentally naïve monkeys and tested them on only four problems; sophisticated monkeys were tested on a large number of problems in the later "duplications" of their experiment. In order to evaluate the two possible explanations of the different experimental results, naïve monkeys were tested on a large number of problems in the present study.

Gellermann (3) has shown that experimentally naïve rhesus monkeys can solve the double alternation problem with ease. Yet attempts (8, 18) to train experimentally sophisticated monkeys on double alternation have been relatively unsuccessful, possibly because new sets of identical stimuli were introduced after each sequence of four or eight responses. Consequently, the objective of the double alternation study was to determine whether monkeys, after extensive training on visual discrimination problems, could master the double alternation problem if the same stimuli were used throughout training.

* Received in the Editorial Office on April 2, 1958.

¹ This research was supported by Grants M-835 and M-1365 from the National Institute of Health, U. S. Public Health Service.

B. METHOD

1. Subjects

Seventeen adolescent and preadolescent rhesus monkeys were studied. All of the Ss were adapted to the laboratory situation by approximately three months' testing in studies of handedness (16), food preference (15), and social dominance (19), but they had no previous training in discrimination learning prior to the first series of pattern discriminations (*PD-I*). During the five months which intervened between the first and second (*PD-II*) series of pattern discriminations, the Ss were trained on discrimination (11, 14), discrimination-reversal, and conditional discrimination problems, with stereometric objects as the stimuli in all the experiments. Four months elapsed between the second and third series (*PD-III*) of pattern discrimination problems; during this period, the monkeys received 45 days' training on the double alternation problem and served in another study of social behavior (9), but they received no further practice in discrimination learning. (Sickness prevented two monkeys from completing *PD-I*.)

2. Apparatus

The Wisconsin General Test Apparatus (3, Figure 1) was used in both experiments.

3. Stimuli

a. Pattern discrimination experiments. The stimuli were regular geometrical figures cut from colored construction paper and mounted on 3 by 3 in. squares of white cardboard. All of the figures were approximately 3 sq. in. in area; similar stimuli have been described in detail elsewhere (13). The patterns were always presented on a grey three-hole test tray.

b. Double alternation. The stimuli, two identical black rectangles, were presented on a white form board containing two food wells.

4. Procedure

a. Pattern discrimination experiments. The daily testing procedure was the same in all three replications; the monkeys were tested on eight 6-trial discriminations per day. The discrimination problems were preceded by one of four types of pre-discrimination experiences, which are defined with respect to the percentage of reinforcement given the pattern stimulus which is correct (+) in the subsequent discrimination (test) trials: (100) the + object was presented alone over the center food well and rewarded 8 times; (50) both the + and — patterns were presented singly four times each,

and rewarded four times each; (0) the — stimulus was presented alone and rewarded eight times; (C) the control condition, where no pretraining was given. On the pre-discrimination trials the single stimulus was always presented over the central one of the three food wells; on test trials the patterns were placed over the lateral food wells.

The non-correction technique was used throughout. The monkeys were tested seven days a week and at the same time of day, 23 hours after receiving their daily food ration.

In *PD-I*, eight 6-trial discriminations were tested daily for 36 days, a total of 288 problems. Eight problems were tested each day for 18 days in both *PD-II* and *PD-III*; the total number of problems was thus 144 in each of these experiments.

b. Double alternation. Ten sequences of four responses each were presented daily for 45 days. Food was placed under the object on the right on Responses 1 and 2, and under the object on the left on Responses 3 and 4. Thus, all animals were trained on the conventional *RRLL* sequence, without differential visual cues. Approximately 10 seconds elapsed between consecutive responses within a sequence, and a standard delay of 30 sec. was interposed between consecutive sequences.

C. RESULTS

1. Pattern Discrimination Experiments

The results are summarized in Table 1, which shows the percentage of correct responses on Test Trials 2 to 6 as a function of the prediscrimination reinforcement conditions on the three replications. The significance of the differences between reinforcement conditions was determined by Scheffé's test (10). Table 1 shows that during *PD-I* performance under Conditions C, 0, and 50 was significantly inferior to discrimination learning following Condition 100, and that performance was but little better than chance even

TABLE 1
PERCENTAGE OF CORRECT RESPONSES ON TRIALS 2 TO 6 ON PATTERN DISCRIMINATION PROBLEMS

Replication	Reinforcement Condition				Mean
	C	0	50	100	
I	54	54	55 *	57	55
II	67	67 *	72	73	70
III	55 *	59	59 *	62	59

* Indicates that the gap between adjacent percentages is significant at the 1 per cent level of confidence (10).

9.3.51
5746

under the latter condition. Markedly superior performance was observed during *PD-II*. The improvement was general, at least 13 per cent more correct responses being made under every condition. Performance under Conditions 50 and 100 was significantly superior to Conditions C and 0. On this replication only did performance on Trial 1 of the problems reflect the effect of preliminary reinforcement. Following Conditions 0 and 100, Trial 1 scores were significantly less and greater than chance, respectively. On *PD-III*, lower scores were obtained than on *PD-II*, and on this replication only Conditions C and 100 produced significantly inferior and superior discrimination, respectively. (It should be noted that the percentages given in Table 1 are based on more than 5000 responses in *PD-I*, and more than 3000 trials in *PD-II* and *PD-III*. Thus, rather small percentage differences represent large and consistent differences in absolute scores.)

2. Double Alternation

The results of the double alternation experiment are summarized in Table 2, which shows the percentage of correct responses and of correct sequences

TABLE 2
PER CENT ERRORS AND CORRECT SEQUENCES IN DOUBLE ALTERNATION LEARNING BY MONKEYS

Sequences	Response				Mean	Correct sequences
	1	2	3	4		
1-150	41	28	54	15	34.5	14
151-300	17	35	60	12	31.0	17
301-450	12	38	48	12	27.5	15

(*RRLL*) on successive thirds of testing. It is readily apparent that learning by the whole group was imperfect within the period of testing. Examination of the individual monkeys' records indicated that 9 of the 17 *Ss* attained the criterion of an average of 80 per cent correct responses over five days.

D. DISCUSSION

The results of the first pattern discrimination experiment indicate that naïve rhesus monkeys are much less capable of discriminating patterned forms than colored patterns (1); this finding is compatible with the observation that highly trained monkeys made more errors in discriminating patterned forms than colors.

In spite of the very low level of success attained by the *Ss* in *PD-I*, a small but statistically significant superiority of performance was observed

under the reinforcement Condition 100, which permitted positive transfer from the single to the paired stimulus situation. On all three replications discrimination performance following reinforcement Condition 100 was significantly superior to performance following preliminary training calculated to produce negative transfer (0) or no transfer (C). Thus, the results of the pattern discrimination experiments again fail to confirm the results of Lashley and Wade, and lend some support to the continuity theory of discrimination learning (12).

The results of the double alternation experiment indicate that monkeys which have had extensive training on visual discrimination problems learn double alternation much more slowly than experimentally naïve monkeys, even when the stimuli are constant throughout training. Consequently, the poor learning of double alternation previously noted (8, 18) in highly trained monkeys cannot be attributed to the nature of the manipulanda.

The fundamental question raised by these observations is the transfer of learning sets. The monkeys' overall performance was 15 per cent better on *PD-II* than on the first series of pattern discriminations, but following tuition on double alternation their performance on *PD-III* declined 11 per cent from the high point reached on *PD-II*. Also, it was observed that after several months of training on discrimination problems, the monkeys were less competent on double alternation than experimentally naïve animals.

At present one can do little more than to suggest an *ad hoc* hypothesis to account for these findings. Efficient discrimination of patterned forms is probably dependent upon the monkeys' acquisition of some kind of receptor adjustments or orienting responses (20). But the naïve animal can not acquire these responses when initially tested with perceptually difficult stimuli. However, such orienting responses are acquired rapidly during training on the discrimination of objects, which are easier to differentiate than patterns. Once formed, the orienting responses may be applied to the discrimination of more difficult stimuli. In such a fashion one can account for the low level of performance in *PD-I*, and the considerable improvement noted in *PD-II*.

Harlow's (5) analyses have shown that position preferences are virtually eliminated in the formation of learning sets by most monkeys. So it is likely that naïve monkeys have stronger spatial hypotheses than do animals subjected to months of training on problems in which positional cues are irrelevant, and are therefore able to solve a spatially defined problem like double alternation more quickly. Training on double alternation, however, may be assumed to revive the tendencies of sophisticated monkeys to respond

to position, and thereby to disrupt the set to respond to visual rather than spatial differentia, which might account for the regression of performance on PD-III.

The preceding account of the results is obviously speculative, since such large positive and negative transfer effects were not anticipated and appropriate control groups were not provided for in the experimental design. The purpose of this discussion has been to formulate the problem of transfer relations among learning sets. It remains for future research to demonstrate unambiguously the existence of positive and negative transfer of learning sets, and to define the variables which condition such effects. Currently, it is reasonable to suspect that the magnitude of transfer effects is related to task difficulty, since only minor interference (4) or none at all (14) is observed when monkeys are tested intercurrently on simple object and positional discriminations.

E. SUMMARY

A group of 17 rhesus monkeys was tested on three series of patterned form discrimination problems and on the double alternation problem. The animals were experimentally naïve when tested on the first series of 288 pattern problems, received extensive training in discriminating objects between the first and second series (144 problems), and were tested for 45 days on double alternation between the second and third series of 144 pattern discriminations. The following results were obtained:

1. The monkeys averaged 55, 70, and 59 per cent correct responses on the first, second, and third series of pattern discriminations, respectively.
2. The trained monkeys did not solve the double alternation problem within 45 days, and were thus less competent on this task than Gellermann's naïve monkeys.

The results suggested that the formation of specific learning sets may have important transfer effects with respect to the solution of later problems. Positive transfer is indicated by the improvement from the first to second series of pattern discriminations; negative transfer by the inferior performance observed on double alternation and the third series of pattern problems.

REFERENCES

1. BLAZEK, N. C., & HARLOW, H. F. Persistence of performance differences on discriminations of varying difficulty. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 86-89.
2. GRANDINE, L., & HARLOW, H. F. Generalization of the characteristics of a single learned stimulus by monkeys. *J. Comp. & Physiol. Psychol.*, 1948, **41**, 327-338.

3. GELLERMANN, L. W. The double alternation problem: III. The behavior of monkeys in a double alternation box-apparatus. *J. Genet. Psychol.*, 1931, **39**, 359-392.
4. HARLOW, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
5. ———. Analysis of discrimination learning by monkeys. *J. Exp. Psychol.*, 1950, **40**, 26-39.
6. HARLOW, H. F., & POCH, S. Discrimination generalization by macaque monkeys to unidimensional and multidimensional stimuli. *J. Comp. Psychol.*, 1945, **38**, 353-365.
7. LASHLEY, K. S., & WADE, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72-87.
8. LEARY, R. W., HARLOW, H. F., SETTLAGE, P. H., & GREENWOOD, D. D. Performance on double alternation problems by normal and brain-injured monkeys. *J. Comp. & Physiol. Psychol.*, 1952, **43**, 419-429.
9. MARONEY, R. M., WARREN, J. M., & SINHA, M. M. Stability of social dominance hierarchies in monkeys (*Macaca mulatta*). *J. Soc. Psychol.* (in press).
10. SCHEFFÉ, H. A method of judging all contrasts in the analysis of variance. *Biometrika*, 1953, **40**, 87-104.
11. SINHA, M. M. Effects of reinforcing single stimuli upon subsequent discrimination learning by rhesus monkeys. *J. Comp. & Physiol. Psychol.* (in press).
12. SPENCE, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427-449.
13. WARREN, J. M. Additivity of cues in visual pattern discrimination by monkeys. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 484-486.
14. ———. Solution of object and positional discriminations by rhesus monkeys. *J. Comp. & Physiol. Psychol.* (in press).
15. ———. Effect of satiation on food preferences in monkeys. *J. Genet. Psychol.* (in press).
16. ———. The development of paw preferences in cats and monkeys. *J. Genet. Psychol.* (in press).
17. WARREN, J. M., & HALL, J. F. Discrimination of visual patterns as a function of motivation and frequency of reinforcement. *J. Genet. Psychol.*, 1956, **88**, 245-249.
18. WARREN, J. M., LEARY, R. W., HARLOW, H. F., & FRENCH, G. M. Function of association cortex in monkeys. *Brit. J. Anim. Behav.*, 1957, **5**, 131-138.
19. WARREN, J. M., & MARONEY, R. M. Competitive social interaction between monkeys. *J. Soc. Psychol.* (in press).
20. WYCKOFF, L. B., JR. The rôle of observing responses in discrimination learning: Part I. *Psychol. Rev.*, 1952, **59**, 431-442.

Yerkes Laboratories
 Orange Park
 Florida

SOME ASPECTS OF THE BEHAVIOR OF THE ANT IN THE MAZE UNDER VARIOUS CONDITIONS OF STIMULATION*

Department of Psychology, The George Washington University

WILLARD E. CALDWELL AND MINNIE ROB PHAUP

A. PROBLEM

The European Ethologists, Tinbergen (7) and Lorenz (+), have raised the instinct problem again, and this has served in a number of ways as a stimulus for American comparative psychologists. First, it has encouraged the study of species other than the albino rat or the rhesus monkey. Secondly, it has suggested that types of drives other than hunger and thirst be studied. Thirdly, it has prompted American psychologists to try to observe more of the behavior that is going on in the laboratory, and particularly that which does not fit in with the conventional theories.

Though the proverbial command, "Go to the ant . . . consider his ways," was originally addressed to sluggards rather than to scientists, the ant has been selected as the subject for this study. It is a suitable organism for investigation and has been studied in the laboratory by such comparative psychologists as Schneirla (6). There is also a considerable body of data upon the ant both experimental and anecdotal.

The present study is a continuation of a series of experiments performed by Caldwell and associates, and is derived from a field theory outlined by Caldwell (1 and 2). In one experiment, performed by Cole and Caldwell (3), goldfish were subjected to a brightly lighted maze and were allowed to swim into a dark goal box for reward. Time, error, and goal curves were obtained. The present study of the ant is a continuation of the series.

This investigation was designed to (a) determine goal, time, and error curves with ultra-violet and bright light as motivation and reward; (b) observe the general behavior of the ants under a number of experimental conditions such as removal of antennae, varying intensities of light, etc.

B. APPARATUS

The maze used in this experiment was of transparent plastic. The exterior of both the bottom and the sides was painted with white plastic paint to increase the reflective power inside without marring the smooth surface

* Received in the Editorial Office on April 2, 1958.

of the interior. The maze, eight inches long, seven inches wide, and two-thirds of an inch high, contained four blind alleys two inches long and one inch wide (see Fig. 1). The ants lived in small plastic boxes also painted white on the outside. These transport boxes, one inch long, one inch wide, and approximately two-thirds of an inch tall, could be placed in the maze at the entrance compartment and could also serve as goal boxes at the end

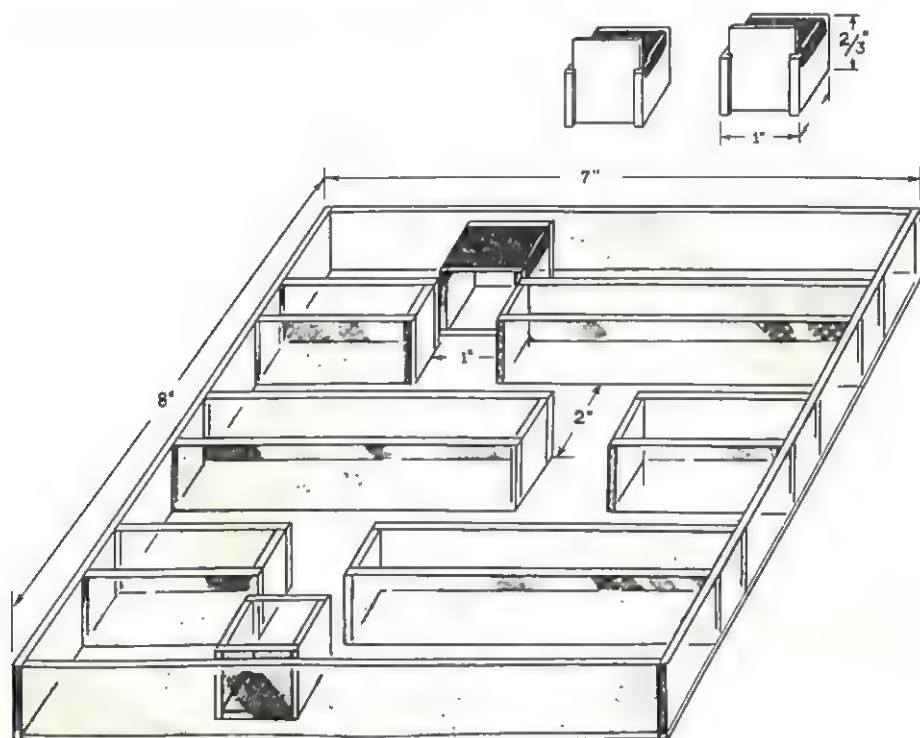


FIGURE 1
DIMENSIONS OF THE MAZE AND PLASTIC BOXES

of the maze. The entire maze was covered with a removable plastic top which contained four small slot-like openings, the first two to permit the experimenter to open and close the entrance compartment and the goal compartment, the third to serve as a retrace gate dividing the maze in half, the fourth allowing for the blocking of the entrance compartment after the ant had entered the maze. Both the retrace gates and the doors to the transport boxes were one and one-half inches high, thus allowing the experimenter to manipulate them above the top of the maze. The slots served the additional purpose of letting air into the maze while the lid was down.

A mineral light model *L* ultraviolet lamp, long wave (3660 Angstrom units wave length), and operating on 110 volt *AC*, 50-60 cycle, was utilized in the experiment since its light was not so intense that it would injure the ants or the human eye and also because of the negligible amount of heat emitted. The lamp was provided with a transformer and a tube-like shade which was suspended about six inches above the middle of the maze so that the light could be diffused equally throughout the maze.

In another part of the experiment a seven and one-half watt light bulb enclosed in a small desk lamp and located about six inches above the center of the maze provided light whose mean intensity was approximately three foot-candles, as measured with an exposure meter. If more intense light was required the seven and one-half watt bulb was replaced by a 60-watt bulb, thus insuring a median reading of 85 foot-candles when the light was measured by the exposure meter. The exposure meter used was a Weston Master II Universal, model 735 (Western Electric Inst. Corp.). Because of slight shadows in the maze, it was necessary to measure the light intensity from several different places, the median intensity serving as the indicator. The room in which the experiment was performed was totally dark; its walls and the table upon which the maze rested were painted black in order to keep the light more uniform.

The transport boxes could be completely darkened by covering the tops with black plastic tape which could be removed according to the dictates of the experiment. When the boxes were covered with the black tape, the intensity of light in them was less than .2 foot-candles.

Temperature was measured by two small thermometers placed inside the maze. Daily measurements of humidity in the laboratory were also taken.

C. SUBJECTS

Prior to the initiation of the experiment a pilot study was conducted to test the ants' reactions to living conditions of the laboratory, to the apparatus, and to the plastic transport boxes. For this preliminary study 100 carpenter ants (workers), obtained from the Carolina Biological Supply Company, Elon College, North Carolina, were used. They were kept in a jar filled with natural material from the ants' native habitat. A moss pad on top of the jar provided a place on which the ants could deposit their wastes and also preserved the proper humidity. It was important, however, not to allow the pad to become too wet.

The ants' food was a block of sugar placed on top of the moss pad when the previous food had been consumed. After 10 days in the jar 50 of the

ants were transferred to individual plastic transport boxes, where they remained for one month. During this month the experimenter noted the death rate, effect of different amounts of water being placed in the boxes, effects of different kinds and amounts of food placed in the boxes, the best way of handling the ants, and their general reactions to being put in the maze as well as to various intensities of light.

Results of this month-long pilot study indicated (*a*) a small droplet of water daily kept the ant alive and healthy; when more than this was introduced the ant usually died; (*b*) two or three crumbs of toast kept away from the water and changed daily to prevent mildewing obtained optimum survival.

Upon completion of the pilot study 100 additional ants of the same kind were obtained, 36 of which were divided into three groups of 12 and given two days in individual plastic boxes before the experiment proper was begun.

D. PROCEDURE

In Part I of this experiment, an attempt was made to investigate the effect of the exploratory drive on the maze behavior of the ants. Twelve ants were run for 19 trials. The ants lived in individual plastic boxes. For 15 minutes before each run a light of the same intensity as that in the maze was turned on over their boxes. The amount of light reflected from the boxes was approximately three foot-candles. Then the ants in their transport boxes were placed very carefully in the maze. The light in the maze and in the goal box was the same, three foot-candles. After the transport box was introduced into the maze, the experimenter raised the door of the box and allowed the ant to enter the maze. Runs were made once a day, usually between one and five P.M. The ant was given five minutes to reach the goal box, at the end of which time, if the ant had not succeeded, it was placed in the goal box by the experimenter. Should the ant reach the goal box in less than five minutes, the time (in seconds) for that trial was recorded. If the ant went all of the way into any of the four blind alleys, an error was noted. Retraces were not counted; a retrace gate separating Blind Alleys 1 and 2 from 3 and 4 was inserted. At the end of the runs, the ants were taken back in their goal boxes and allowed to remain in the light for 15 minutes. The above procedures were adopted to ascertain the motivating effects of the stimulus properties of the maze itself, the light in the apparatus playing as small a rôle as possible. This group became the control group and served as a reference point for comparison of the other groups.

In Part II of the experiment, the procedure was essentially the same as in Part I, except that the ants were kept in plastic boxes covered with black plastic tape. As soon as the box was placed in the maze, the tape was removed. The goal box, however, which served also as living quarters for the ant until the next trial, was covered with the black tape. Motivation for the 19 trials for this group consisted of the presentation of ultraviolet light.

Part III was identical to Part II with the exception of the use of intense light, measuring 80 foot-candles as motivation.

In the final Part IV of the experiment five ants from the exploratory group and five from the ultraviolet group had their antennae clipped very close to their heads, with a small pair of clippers. These ants were run for five trials in accordance with the procedure used on the exploratory group. Their performances were then compared with their performances on the last five of their 19 trials.

Odor presented an important variable to control in this entire experiment, particularly in view of the trailing reaction of the ant. Therefore the maze and the goal boxes were not cleaned with any chemical but were sponged thoroughly and dried with soft tissue, the maze after each trial and the goal box before it was placed in the maze. This process tended also to keep down the accumulated heat.

In addition to quantitative observations, the experimenter made observations of a qualitative nature on both the pilot and the experimental ants while they were in the jar, when they were in the plastic boxes and in the maze, and finally while they were in a large fishbowl containing only sand.

E. RESULTS AND DISCUSSION

An analysis of variance was run on the error data for the ultraviolet, bright light, and exploratory groups for Trials 1-5, 6-10, 11-15, and 16-19. For Trials 1-5 there was an F of 10.02, significant at the one per cent level. An F of 2.5 for Trials 6-10 was not significant. The variance of the bright light group, 438.8, was out of proportion to the others. This proves interesting since this group received intense stimulation and thus exhibited more random behavior than did the others. For Trials 11-15 the F of 3.31 was significant at the five per cent level while for Trials 16-19 the F of 15.79 was significant at the one per cent level.

Scheffe (5) has devised a method for judging all contrasts in the analysis of variance. Although analysis of variance may not really be valid for frequencies, it appears that no assumptions are seriously violated.

For Trials 1-5 the contrast between the ultraviolet and the bright light groups and that between the bright light and the exploratory groups was significant at the five per cent level (the level to be used henceforth) while the contrast between the ultraviolet and the exploratory groups was not significant.

Trials 6-10 showed no significant contrasts; in Trials 11-15 the contrast between the ultraviolet and the bright light group was significant, whereas contrasts between the ultraviolet and the exploratory groups, and the bright light and exploratory groups were not significant.

Trials 16-19 offered significant contrasts between the ultraviolet and the bright light groups and between the ultraviolet and exploratory groups, but no significant contrast appeared between the bright light and the exploratory groups.

Analysis of variance was not applied to the time data because the distributions appeared too skewed to be normal even under usual transformation. The skewness was ascribed to truncation at five minutes.

Contingency chi-square, applied to the data representing the number of ants who reached the goal box within the allotted five minutes showed no significant differences among the three groups.

One of the indices used in comparing the performances of the ants in the three parts of the experiment was the percentage of ants reaching the goal box within the time limit. This was obtained for each of the 19 trials and results are presented in Figure 2.

The group receiving the 80 foot-candle light appeared highly motivated from the start (especially in Trials 1-8) in comparison to the other two groups. The other two groups, particularly the ultraviolet group, seemed inhibited at first, tending to remain immobile in the entrance compartment. Up to Trial 10, the curve for the ultraviolet group ascends much more gradually than do those for the others.

Since a consideration of individual differences is important, especially in the light of the common tendency to think of ants as exhibiting predominantly an "instinctive" type of responses, the presentation of individual learning curves seems desirable. Figure 3 contains the time curves of eight of the ants from the ultraviolet group, the variability among these being rather typical of that exhibited in other parts of the experiment.

The curves of Ants 2, 3, and 7 strikingly resemble the more conventional learning curves while the performances of Ants 11 and 12 show almost no signs of learning. It will be noticed that most of the curves begin to rise

after Trial 12. In general, the ants exhibited variability from each other, from trial to trial, and within a single trial.

Figure 4 presents the mean time in seconds for the three parts of the experiment. These three curves show a tendency to descend in somewhat the same way as a general learning curve does; then around Trial 11 or 12 they begin to ascend, that for the exploratory group showing a very rapid

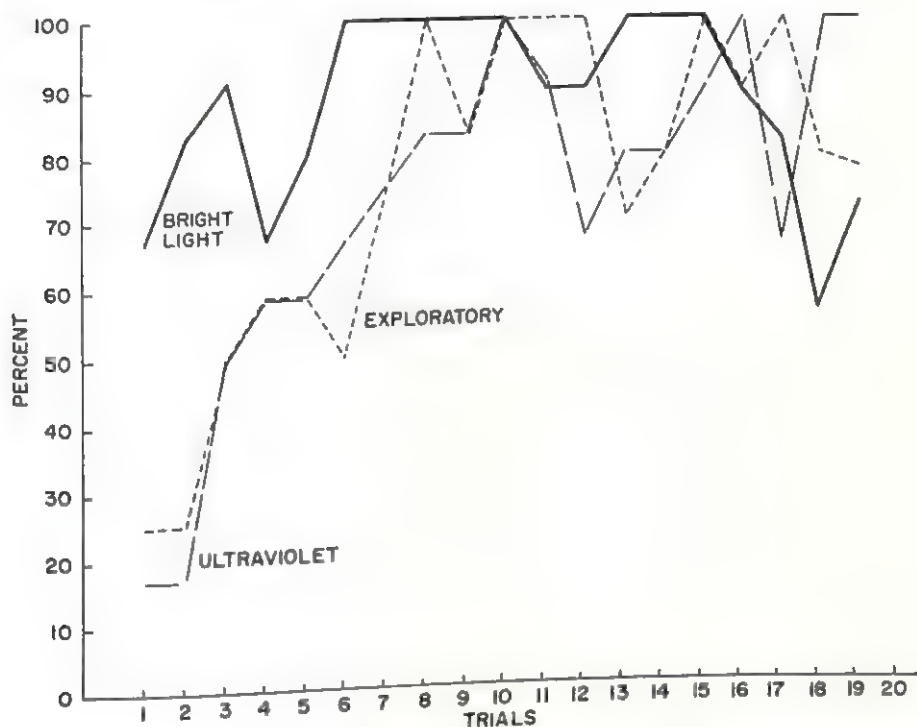


FIGURE 2
PERCENTAGE OF ANTS REACHING THE GOAL BOX WITHIN THE FIVE-MINUTE TIME LIMIT, FOR TRIALS 1-20

rise. The bright light group shows a much more gradual increase, and especially in Trials 1-15 there is a striking and consistent difference between this group and the other two.

Figure 5 reports the error data from the three groups for the 19 trials. Both the exploratory and the bright light groups show a gradual increase in errors throughout the 19 trials, but the ultraviolet group exhibits a gradual increase only until Trial 6 when the number of errors begins to decrease in a manner similar to that of the conventional error curve. Except in

Trial 11, the bright light group made more errors for all trials than did the ultraviolet group. It is interesting to notice that on Trial 11 the ultraviolet group begins to show an error increase while the incidence of error for the other two groups begins to decrease.

Error and time data for the ants with clipped antennae are presented in Table 1. Their pre-clipped performances on Trials 15-19 were compared with the performances after the antennae were clipped. Most of these ants

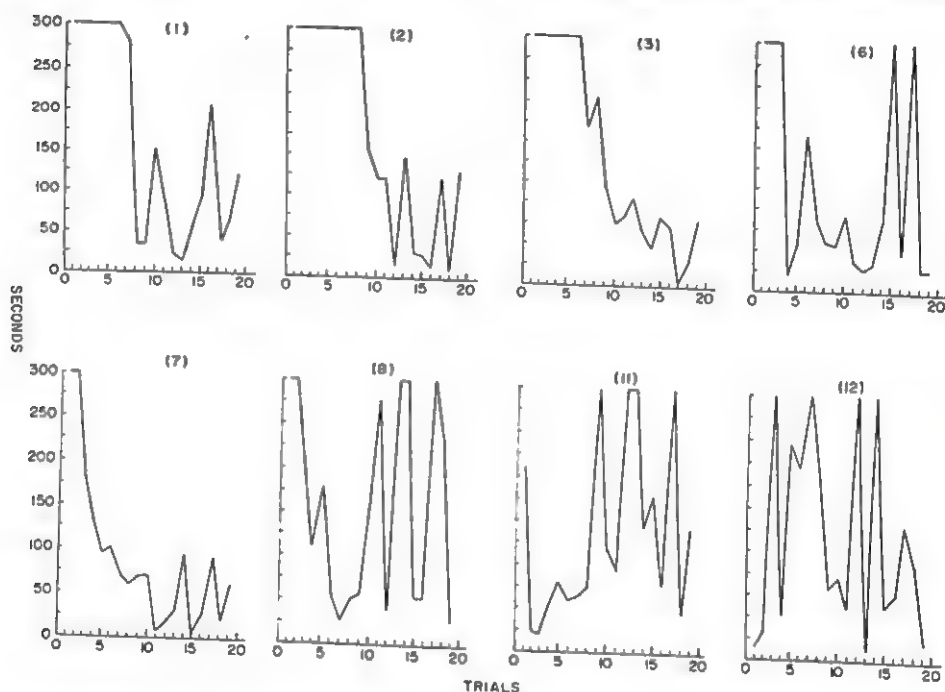


FIGURE 3
TIME CURVES FOR INDIVIDUAL ANTS ILLUSTRATING VARIABILITY IN PERFORMANCE

had originally reached the goal box in less than five minutes; after the antennae were removed only Ants 11 and 2 reached it—and they only twice. Three of the clipped ants died after the second trial. Some of the antennaeless ants were practically immobilized from the beginning; others became immobilized later, and some continued active even though they seldom reached the goal box—again strikingly demonstrating the principle of individual differences. Furthermore, these ants “indulged” in some interesting behavior patterns in the maze. They would crawl through the maze sideways, using their legs as though they were the missing antennae or moving their heads back and forth as they had previously done their antennae.

Throughout all the phases of the experiment theoretically provocative qualitative observations of the behavior of the ants were made. One recurring behavior pattern was the stroking of the antennae with the legs. Also the *E* observed what might be called a tracing response; some ants would go over the same pathway again and again when they were in the maze, the

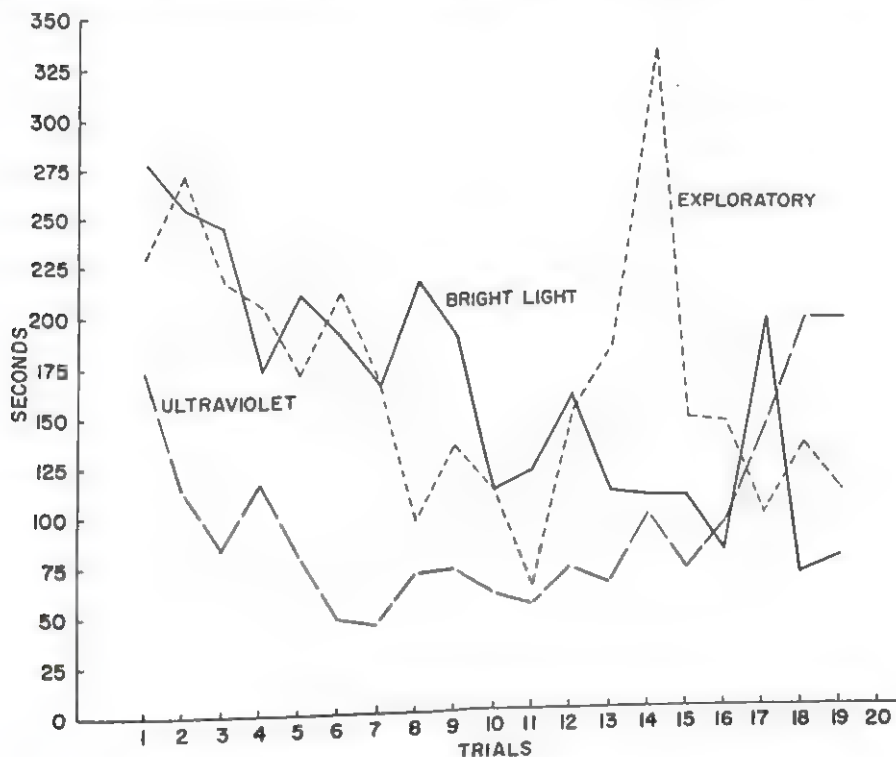


FIGURE 4
MEAN TIME IN SECONDS FOR THE ULTRAVIOLET, BRIGHT LIGHT, AND EXPLORATORY GROUPS
FOR TRIALS 1-20

pathway varying from ant to ant and from trial to trial. Since the maze was cleaned after each trial, this suggested that the ants were retracing their previous trails.

The pilot ants in the jar in which they were living before the experiment began, tended to cluster together when the jar was chilled or tapped. The ants exhibited too a rather precise reflex to gravitational changes. This manifested itself when the *E* would turn a net upon which the ants were crawling and the ants almost invariably righted themselves.

The *E* placed 50 of the ants in a goldfish bowl which had sand in the bottom and a cone-shaped piece of paper in the center. When the paper was shaken, the ants ran out but in a few seconds crawled back into the cone. If the paper was removed, the ants would remain together, almost motionless. If they were disbanded, they would quickly join each other again.

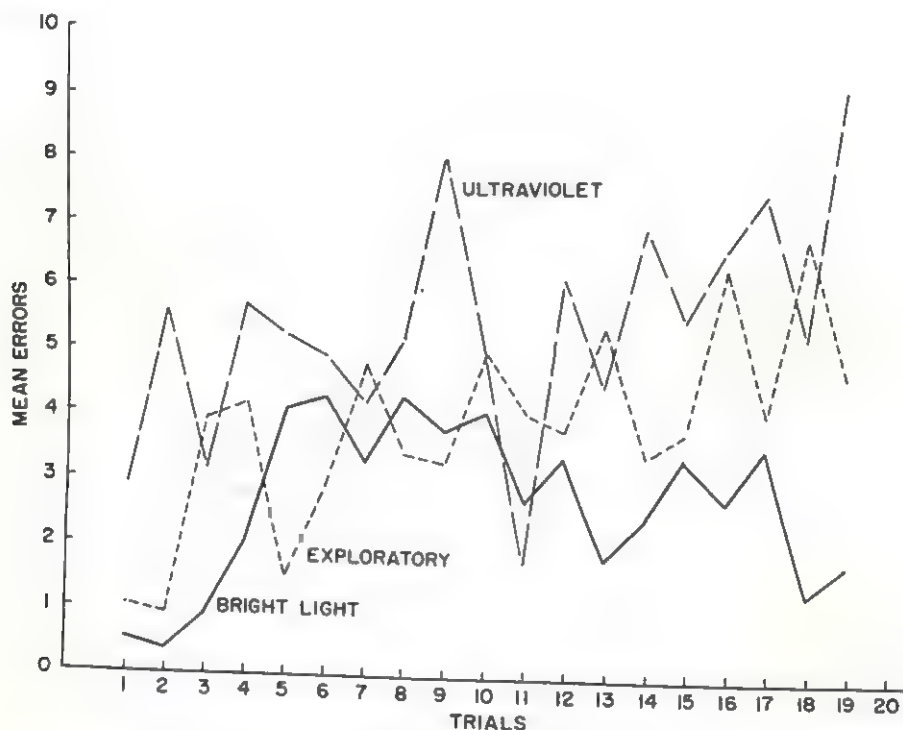


FIGURE 5
MEAN ERRORS FOR THE ULTRAVIOLET, EXPLORATORY, AND BRIGHT LIGHT GROUPS FOR TRIALS 1-20

F. SUMMARY AND CONCLUSIONS

Three groups of 12 ants each were given 19 trials in a maze. The exploratory group were subjected while in the maze and for 15 minutes before entering it and subsequent to removal from it to illumination measuring three foot-candles. The bright light group were kept in the dark except in the maze which had an intensity of 80 foot-candles though the goal box was dark, the light in it measuring less than .2 foot-candles. The third group were also kept in the dark, except that in the maze these ants were

subjected to an ultraviolet light. The goal box for this group too was dark.

At the end of the 19 trials five ants from the exploratory group and five from the ultraviolet group had their antennae clipped and were then given five trials in the maze with the same procedure as that used for the exploratory group.

TABLE 1
TIME AND ERROR DATA FOR ANTS BEFORE AND AFTER ANTENNAE CLIPPING

Trial	Exploratory Ants						Ultraviolet Ants				
	1	2	3	4	10	11	1	2	3	8	12
<i>Time</i>											
15	214	129	20	35	60	120	92	27	83	53	180
16	55	101	27	160	185	300	205	14	70	152	75
17	26	41	21	10	180	180	39	120	6	300	155
18	30	300	167	156	120	85	65	10	30	235	110
19	12	30	15	300	7	300	120	131	80	27	20
1	300	140	300	300	300	275	300	300	300	300	300
2	300	300	300	300	300	300	300	300	300	300	300
3	300	300	—	—	300	300	—	300	300	300	300
4	300	300	—	—	300	300	—	300	300	300	300
5	300	188	—	—	300	300	—	300	300	300	300
<i>Errors</i>											
15	4	6	1	3	3	5	3	1	3	2	2
16	2	8	4	16	5	12	6	0	0	3	1
17	2	0	3	1	8	8	5	5	0	2	1
18	4	9	9	12	5	6	1	0	4	2	0
19	3	2	2	6	2	11	4	3	3	2	0
1	0	4	0	3	8	0	0	0	3	0	0
2	5	0	0	0	1	3	1	1	1	0	0
3	0	0	—	—	0	3	—	0	0	6	0
4	0	3	—	—	0	6	—	1	0	0	0
5	0	4	—	—	2	0	—	8	0	1	0

A five-minute time limit was imposed on all of the trials. Time, errors, and percentage of ants to reach the goal box were obtained as indices. The ants lived in plastic boxes which enabled the *E* to transport them to the maze and back and which also served as entrance and as goal boxes.

Data from the experiment suggest the following conclusions:

1. An analysis of variance run on the error data for the ultraviolet, bright light, and exploratory groups for Trials 1-5, 6-10, 11-15, and 16-19 showed that for Trials 1-5 the *F* of 10.02 was significant at the one per cent level; for Trials 6-10 the *F* of 2.5 was not significant; for Trials 11-15 the *F* of 3.31 was significant at the five per cent level; and for Trials 16-19 the *F* of 15.79 was significant at the one per cent level.

2. When Scheffe's (5) method for judging all contrasts in the analysis

of variance was used, the contrast between the ultraviolet and the bright light groups was significant at the five per cent level as was the contrast between the bright light and the exploratory groups. For Trials 6-10 none of the contrasts was significant, but for Trials 11-15 and 16-19 the contrasts between the ultraviolet and the bright light groups were both significant. For Trials 16-19 the contrast between the ultraviolet and the exploratory groups was also significant.

3. Since the distributions were too skewed to be normal even under usual transformation, analysis of variance was not applied to the time data. This was due to truncation at five minutes.

4. Contingency chi-square applied to the data representing number of ants who reached the goal box within the time limit revealed no significant difference among the three groups.

5. Marked individual variability in the ants' performances manifested itself in regard to both learning and motivation. Some ants demonstrated conventional learning curves for both time and errors while others exhibited rather random performances. If errors may be used as one index of activity level, or motivation, there were notable individual differences among the ants in this regard. Some were extremely active; others almost inactive in the maze.

6. The three different types of situations seemed to produce different responses in the ants. Activity in the very bright light situation was greater than that in the other two situations. This was particularly true in the beginning of the experiment when the ultraviolet light appeared to act as an inhibitor. However, the ultraviolet group showed more goal-oriented behavior toward the dark box.

7. When the antennae were clipped, some of the ants died, some of them remained relatively immobile, others explored the maze. Some tended to substitute their heads for their antennae; some thrust their legs forward as if they were antennae.

8. Qualitative observation made on the 36 ants used in the experiment and on the 100 pilot ants remarked the following phenomena: if the jar in which some of the ants lived were tapped or chilled the ants would descend to the bottom; some crawled under a paper cone placed in the bottom of a goldfish bowl; and they were exceedingly sensitive to gravitational changes.

9. The techniques used in this experiment may be of value as a standardized approach to the study of learning problems and of other problems pertaining to genetically determined behavior.

REFERENCES

1. CALDWELL, W. E. The mathematical formulation of a unified field theory. *Psychol. Rev.*, 1953, **60**, 64-72.
2. ———. Field theory: II. Some mathematical applications to comparative psychology. *Psychol. Rev.*, 1954, **61**, 271-275.
3. COLE, M. B., & CALDWELL, W. E. The utilization of light as exteroceptive motivation in the comet goldfish (*Carassius auratus*). *J. Comp. & Physiol. Psychol.*, 1956, **49**, 71-76.
4. LORENZ, K. Z. The comparative method in studying innate behavior patterns. In *Symposia of the Society for Experimental Biology*. No. 4. *Physiological mechanisms in animal behavior*. New York: Academic Press, 1950.
5. SCHEFFE, H. A method for judging all contrasts in the analysis of variance. *Biometrika*, 1953, **40**, 87.
6. SCHNEIRLA, T. C. Learning and orientation in ants. *Comp. Psychol. Monog.*, 1929, No. 30.
7. TINBERGEN, N. *The Study of Instinct*. London: Oxford Univ. Press, 1951.

Department of Psychology
The George Washington University
Washington 6, D.C.

DIFFERENTIAL DEVELOPMENT OF INTELLIGENCE IN THE COLLEGE YEARS*

Department of Psychology, Iowa State University

DON C. CHARLES AND SALLY ANN PRITCHARD

A. PROBLEM

The purpose of this study was to investigate the changes in intelligence test scores which occur through the college years, especially in regard to sex differences. Many studies of intelligence in the college years have been made: representative reports include those by McConnell (5), Sister M. F. Louise (4), Shuey (6), and Hunter (3). There is general agreement in these studies that students do improve their scores during the college years, that those in the lowest levels improve the most, and that most of the gain occurs in the first year. It was hypothesized in this study that the men would gain more than the women, because of the more rapid development of girls up to adolescence (1). The studies of McConnell (5) and Flory (2) considered sex differences but were inconclusive. None of these studies attempted any control for regression or for initial differences of scores between sex groups, both of which must be recognized as critical factors in longitudinal studies.

Two questions were to be answered by this study, then. They were: (a) what changes in intelligence test scores occur during the college years, and (b) is the change related to the sex of the students?

In this study the groups were not equivalent on initial testing and the correlations between initial and final scores were $+.59$ and $+.51$ for men and for women respectively. Analysis of covariance was used to evaluate the data because it not only took into account these initial differences in scores and the degree of correlation between initial and final scores, but also reduced the sampling error.

B. PROCEDURE

1. *Test*

The 1952 edition of the American Council on Education Psychological Examination (A.C.E.) was used for both original and terminal testing. Only total raw scores were used in evaluation.

* Received in the Editorial Office on April 16, 1958.

2. Subjects

A random sample of 112 (61 men and 51 women) Iowa State College students were retested in their fourth year of college. All of these had the same amount of consecutive college education at the same school.

3. Statistical Method

The significance of the amount of change in score for the total group was tested by a Student's *t*-test. Statistical significance of the difference between sexes in amount of change in total raw score was tested by analysis of covariance.

C. RESULTS

The initial, final, and difference scores for men and for women are set forth in Table 1. It will be observed that a total mean increase of +13.55

TABLE 1
INITIAL, FINAL AND DIFFERENCE A.C.E. SCORES

Group	N	\bar{X} initial test	\bar{X} final test	Difference
Men	61	129.56	141.49	11.93
Women	51	119.37	134.84	15.47
Total	112	124.91	138.46	13.55

points occurred. A Student's *t*-test with paired observations showed this increase to be highly significant ($p < .01$). Fifty-three of the 61 men who were retested gained in score on the retest while eight had lower scores. The range of differences between the two testings was from -31 raw score points to +34. Of the 51 women retested, 48 gained in score while three lost. The range of differences of their scores was from -7 to +38. The mean difference score for men was +11.93 and for women it was +15.47.

In this analysis the final mean scores were adjusted for initial differences from 141.49 to 139.54 and from 134.84 to 137.17 raw score points, for men and women respectively. It can be seen that although the difference between the groups decreased, the men still had the higher mean scores.

The covariance analysis showed no significant difference in amount of change in score between men and women. The obtained *F* was 3.48 where an *F* of 3.94 was needed for significance at the .05 level with 1 and 109 degrees of freedom. Thus, the null hypothesis that there was no significant sex difference in change of score was tenable.

D. DISCUSSION

A major problem in longitudinal studies is sample shrinkage and sample representativeness. An accompanying problem is that of regression in retest scores. The subjects of this study were all seniors and this inevitably produced a higher mean score than the mean score of their class as entering freshmen, because of the dropping out of the less able. The senior sample mean, however, was just 6.91 points higher than that of the original freshman norm group, and there was a range of 79 points on the test. It was not likely then that a large part of the change in scores in this study was due to regression toward the mean since the scores of the persons selected were neither unusually high nor low.

As mentioned earlier, only two previous studies seem to have considered sex differences at the college age (2, 5) and both of these used critical ratio to test the significance of the difference between the mean gains of men and women. A technique such as critical ratio or simple *t* test implies a 1-to-1 relationship between initial and final scores, while the covariance adjustment technique used in this study provides a way of correcting final means for initial differences and allows for the degree of correlation between initial and final scores. In the present study and in the two earlier ones, there was a difference in initial test score of the sexes, but in the earlier studies no effort was made to adjust for this difference. As in this study, no differences were found between the sexes in amount of change in score in the studies of Flory and of McConnell, but without recalculation of the data it cannot be determined what would have been found had those authors controlled for the initial differences in scores.

The hypothesis was made in this study that men would show significantly greater change in score on the A.C.E. than women. One of the reasons for this assumption was the fairly general idea that intelligence continues to grow until adolescence and beyond. Another concerns the numerous childhood studies of sex differences in growth of intelligence. Although these studies are not in complete agreement, a majority of them indicate that girls average slightly higher than boys until at least the age of 16 (1, p. 162). Numerous studies of adult intelligence reveal no sex differences in level of ability. It would appear, then, that boys reach the intelligence level of girls sometime between the age of 16 and maturity. Since the men of this study did not make a greater improvement at the .05 level of significance (in fact, the women made an improvement significant at the .10 level) it cannot be assumed that this "catching-up" occurs during the college years of 18-22,

at least for this population. The equalization of the sexes in ability must occur in middle to late adolescence, rather than in early maturity. A much larger population sample in a longitudinal study would appear to be highly desirable to clarify this problem.

E. SUMMARY

The purpose of this study was to answer two questions: (a) Does A.C.E. score change significantly during the college years? (b) Is change in A.C.E. score during the college years related to the sex of the students?

A sample of 112 college seniors (61 men and 51 women) was retested after four years with the A.C.E. The results:

1. There was a significant rise in score on the A.C.E. in the college years for the 112 subjects.

2. An analysis of covariance comparing the final mean scores of men and women adjusted for initial differences showed no significant difference at the .05 level in amount of change in score on the A.C.E. At the .10 level there was a significant sex difference which showed that women gained more.

It was inferred, therefore, that men must make up any inferiority of intelligence test score before college age.

REFERENCES

1. CONRAD, H. S., JONES, H. E., & HSIAO, H., H. Sex differences in mental growth and decline. *J. Educ. Psychol.*, 1933, **24**, 161-169.
2. FLORY, C. D. The intellectual growth of college students. *J. Educ. Res.*, 1942, **33**, 443-451.
3. HUNTER, E. C. Changes in scores of college students on the American Council on Education Psychological Examination at yearly intervals during the college course. *J. Educ. Res.*, 1943, **36**, 284-291.
4. LOUISE, SISTER M. F. Mental growth and development at the college level. *J. Educ. Psychol.*, 1947, **38**, 65-82.
5. MCCONNELL, T. R. Change in scores on the psychological examination of the American Council on Education from freshman to senior year. *J. Educ. Psychol.*, 1934, **25**, 66-69.
6. SHUEY, A. M. Improvement in scores on the American Council on Education Psychological Examination from freshman to senior year. *J. Educ. Psychol.*, 1948, **39**, 417-426.

Department of Psychology
Iowa State University
Ames, Iowa

DISCIPLINE IN THE HOME: A CONTROLLED OBSERVATIONAL STUDY OF PARENTAL PRACTICES*¹

Institute of Child Welfare, University of Minnesota

EDWARD CLIFFORD^{2,3}

A. THE PROBLEM

Research related to home disciplinary practices has taken three general forms: information obtained from children or parents about past parental actions, assessments of home atmospheres, and observations of specific parent-child behavior in the home. The first form mentioned is mainly nonobservational in nature and bears the underlying assumption that some type of a relation exists between the verbal reports of parents and parental action (5). That is, it is possible to infer the actions of parents from questionnaires or attitude scales. The other two forms, observational techniques, on the other hand, often assume the existence of some type of relationship between parental action and attitude. In this case the attitudes of parents are inferred from observations of parental behavior. Aside from these considerations, there are difficulties inherent in the use of both observational and nonobservational

* Received in the Editorial Office on April 9, 1958.

¹ Abridgement of a dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Minnesota, June, 1957.

² A more detailed form of this paper (or extended version, or material supplementary to this article) has been deposited as Document number 5899 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting \$2.50 for photoprints, or \$1.75 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

³ The author is sincerely grateful to Dr. John E. Anderson, under whose direction this investigation was made and who made available the facilities of the Institute of Child Welfare; to Dr. Elizabeth Fuller for making the facilities of the University Nursery School available; to the directors of the various private nursery schools for making their facilities available; to Mr. John Jarelinek, Principal, for making the facilities of the University Elementary School available; to the Minneapolis Board of Education and the principals of the elementary schools involved for making their facilities available; to the teachers from all the schools mentioned for filling out personality ratings on the individual children involved; to the parents who gave so much time and effort in carrying out the lengthy observations required; to the children who served as unknowing subjects; to Mrs. Miriam Clifford for assistance in the collection of data; and to Drs. Solis L. Kates and Albert E. Goss for helpful suggestions.

This study is affectionately dedicated to my wife, Miriam, whose active aid and support made it all possible.

tional techniques. In the observational techniques used, the presence of an observer may introduce uncontrolled variables which affect the character of the observation, while in nonobservational techniques, particularly those dealing with past parental behavior, difficulties lie in the process of recall and the possible distortions involved in this process. Despite these methodological difficulties, research in this area has had an important rôle in contributing to our understanding of parents and children.

It is possible to eliminate one of the difficulties of observational techniques (the presence of a neutral observer in the home). Goodenough (3) introduced a technique which avoids the errors of the interview technique and closely approximates direct observation in the home by trained observers. Through the use of appropriate instruments she demonstrated that mothers could be used as observers of their children's behavior. Briefly, the technique enables parents to observe the behavior of their children and report it in an objective fashion. Moreover, the data are subject to statistical analysis, and the variables under investigation can be controlled through the use of proper forms and directions. With emphasis on formalized description and the avoidance of subjective and evaluative materials, it is possible to systematically investigate the variables that enter into the disciplinary situation. With some modification, this is essentially the technique used in this study.

Using mothers as observers of their own behavior, this study was designed to obtain answers to the following questions.

1. What are the conditions of social interaction which cause discipline? Are there specific areas of social interaction which are more likely to be associated with discipline? An examination of the actions of parents and children in discipline should yield further information about the parent-child relationship in terms of dynamic causation of behavior.

2. What is the relationship between general external conditions and the frequency with which children are involved in disciplinary situations? Are there events, not part of the social interaction of parent and child which are influential in determining the rate of discipline in the home? If such variables can be isolated and a relationship established between them and the rate of discipline, then important practical consequences exist. It would mean that the environment of the child could be manipulated so that the rate of disciplinary problems could be minimized.

3. How are children controlled when disciplinary situations arise and what are the effects of discipline on the child? The effects that are taken into account are those that deal with immediate behavioral correlates and effects in terms of the overall adjustment of the child.

4. Do the differing rôles of parent and child in the disciplinary situation vary with characteristics of the child (age, sex, and adjustment ratings) and with attributes of the parents (mother, father, socio-economic status, and attitudes toward child-rearing practices)?

B. EXPERIMENTAL DESIGN

1. *The Sample*

The sample consisted of 120 mothers of children, 60 boys and 60 girls equally divided among three age levels, three, six, and nine years. Representation was equal from Groups I (*SES* I) and II (*SES* II), of the Minnesota Scale for Paternal Occupations (+).

The sample was restricted to parents whose children were in attendance in a nursery school or an elementary school. Originally, an attempt was made to secure three-year-olds attending an all-day nursery school in order to equate the time spent away from home with elementary school groups. The limited number of subjects attending nursery school all day, however, made this impossible to attain. As a result there are 26 three-year-olds who are in attendance in nursery schools for half-day only.

To further equate the groups, the stipulation was made that both parents had to be living together, the mother had to be a housewife—i.e., not working during the hours the child was in the home, and the children had to be without physical handicap or severe personality disturbances.

The children's ages were between 3-0 and 3-11, 6-0 and 6-11, and 9-0 and 9-11 on the initiating date of the interview with the mother.

2. *Procedure*

The names of parents were obtained from various nursery and elementary school rosters. Letters were sent to the parents over the signatures of the school principals explaining something of the study and inviting their co-operation. The letters were followed by telephone calls and interviews were arranged with those who volunteered to participate in the study. Parents were then interviewed, and, if they met the sampling criteria and were still interested, started on their three-week observations. In order to equate the recording time for all groups, contacts were not made nor observations begun during any vacation period lasting more than two days. No observation was started when the expected terminal date fell more than two days into a vacation period. All forms and directions used were subjected to an analysis of reading ease. They were of medium difficulty on the Flesch (1) scale.

a. Definition of discipline. A mimeographed statement handed to parents defined discipline as follows:

In this study discipline is considered an attempt to correct or modify the child's behavior. Discipline shall include only those incidents in which conflict is involved. That is, there is some resistance on the part of the child. A request by the parent which is complied with easily and immediately would *not* be considered discipline. Routine activities to which a "no" or a "don't" is said to the child should *not* be included unless some conflict develops which makes it necessary for the adult to intervene and attempt to resolve it through further verbal or physical action. For example, if a young child reached for an ashtray, was told "no" and he complied, it would not be considered discipline. If, on the other hand, after the child was told "no" he reached for the ashtray anyway, it would be considered a disciplinary incident in this study. If there is any doubt as to whether a situation is disciplinary or not, record it anyway with a short note or question at the top of the page.

Parents were given an opportunity to discuss this definition of discipline, but insistence was upon following the above definition. This definition had been decided upon in order to reduce the amount of recording done by mothers of young children. The attempt to get away from including the number of times a mother says "no" or "don't" to a child was deliberate, since it was felt that this would increase the recording burden and would not add materially to a picture of discipline as practiced in the home.

b. Parent interview. Parents were given a brief mimeographed statement explaining the study and giving the definition of discipline previously described. They were shown the observational forms and the nature of their task was explained. Those who continued with the study were then asked to fill in a questionnaire giving background information about the child and the family. Following this they were asked to estimate how frequently they used various measures of disciplinary control. Attitudes toward various child-rearing practices were then obtained through the use of the Minnesota Scale of Parents' Opinions. Finally the parent filled out an adjustment inventory for the child, the *Personality Profile* developed at the Institute of Child Welfare of the University of Minnesota.

Filling out the forms mentioned, explaining the nature of the work required, and giving the parents the necessary material for the three weeks of continuous observations completed the interview with the mother. The time required for this interview was approximately 45 minutes.

c. Observational material. Two instruments, attendant instructions and aids along with three stamped, addressed envelopes were left with the parents.

Parents were given 21 forms titled *Daily Information Sheet: Discipline* (Appendix A). This instrument was designed to get at various events which might possibly have an effect on the occurrence of discipline. It provides for the collection of information concerning the sleep of the child the previous night, events occurring during the day, the daily health record of the child, the number of hours the child was under observation that day, and whether or not the parents were paying attention to particular problems or circumstances that day. Each day the child was under observation, a new Daily Information Sheet was filled out regardless of whether discipline occurred that day or not.

A number of forms, sufficient to carry parents through the required three-week observational period, titled *Observational Record: Discipline* (Appendix B) were left with the mothers. This instrument makes it possible to obtain consistent observations from a group of parents, provided they can read and interpret the directions correctly. This form enables the parent to describe the disciplinary situation as it arises in the home without becoming involved in justification and rationalization for their actions. The Observational Record calls for a descriptive account of what actually happened, and it is not concerned with the rightness or wrongness, correctness or incorrectness of the actions of the parent. This was further reinforced in both the written and oral directions given to the parent. With the aid of the Observational Record parents recorded the disciplinary incident in terms of the time of occurrence, the place of occurrence, and the duration of the incident. Further recording concerned the difficulty that arose, the methods used to control the situation, the persons using the control, the issue involved, the outcome of the incident, how the child reacted to the use of disciplinary measures, and how long this latter behavior lasted. For each incident a new form was filled out.

To aid the parents in filling out the required forms properly, mimeographed directions were distributed to them. The directions were made as explicit as possible, and the parents were told to consult these directions continually during the period of observation.

An alphabetical list of methods of control was also given to each parent. The 26 items on this list were numbered in order to reduce the work necessary in keeping the observations. If the parent used a measure of control identical to that of one on the list, she then could enter the number of that control in the appropriate space on the observational record. This helped reduce writing and kept fatigue at a minimum. As a further aid, the meanings of the methods of control listed were defined explicitly, and accom-

panied the list of methods. Parents were told that when a measure of control used did not coincide with one on the list, a sentence or two describing what was done should be inserted in the appropriate space on the Observational Record.

d. Adjustment inventory. Within one month after the completion of the observations by the parent, she was asked to fill out Division III (Social adj.) and IV (Personal adj.) of the Haggerty-Olson-Wickman Behavior Rating Schedules hereinafter referred to as the *H-O-W* III and *H-O-W* IV respectively. The parents had been told that they would be asked to fill out such a form sometime after they completed their observations. The forms were mailed to the parent and an addressed, stamped envelope was provided for the return of the filled out form.

e. Teacher materials. Teachers of the individual children were asked to fill out the Personality Profile as soon as the child became a subject in this study. The teachers were also asked to fill out the *H-O-W* III and *H-O-W* IV for each child within one month after the observations on the child had been completed. The directions for filling out the Personality Profile and the *H-O-W* scales were identical to those given to the parents.

C. RESULTS

Data relevant to several areas will be presented. The areas to be considered are the frequency of discipline, extrinsic factors associated with discipline, the content of disciplinary incidents, the methods of discipline, the outcome of discipline, and adjustment and parental attitude.

1. Frequency of Discipline

How frequently are children disciplined? Three-year-olds are involved in disciplinary incidents on the average, approximately once a day (1.01), six-year-old children are involved in such incidents approximately every other day (.51), while the nine-year-old children are involved approximately every fourth day (.28). The differences that arise between the total age groups, are all highly significant ($p = .001$).

Sex differences are significant at the six-year-old level ($p = .01$). SES differences achieved significance at the three- and nine-year-old levels. The overall differences between age groups are greater than the differences that arise due to sex or socioeconomic factors within the group.

a. Time of day. The data were analyzed in terms of hour of occurrence for the 21-day period. These data are given in Figure 1. For the sake of convenience, the time dimension is broken in the following manner: All

incidents occurring between 1:00 a.m. and 7:59 a.m. are labelled "before 8:00 a.m." Then incidents are recorded in hourly intervals, 8:00 a.m., for example, represents the interval 8:00 a.m.-8:59 a.m. The period from 8:00 p.m. to 12:59 a.m. is called "8:00 p.m. and later."

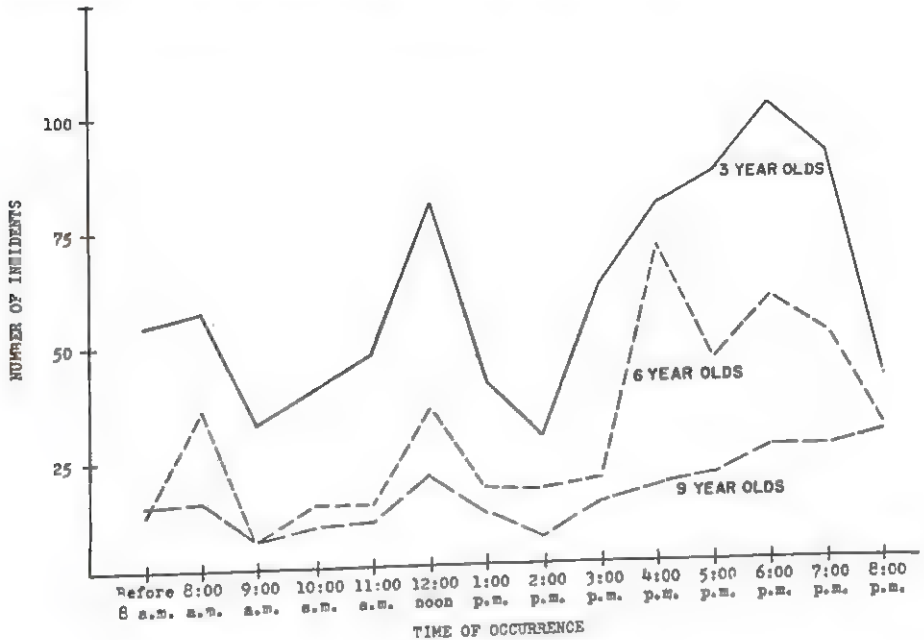


FIGURE 1
THE FREQUENCY OF DISCIPLINE BY THE HOUR OF OCCURRENCE AND AGE

The three-year-olds demonstrate three peak periods during the day when discipline is more likely to occur. These are the 8:00 a.m. hour, the noon hour, and 6:00 p.m. The period of 4:00 p.m. to 7:59 p.m., however, is higher in rate of incidence of discipline than any other period of the day. The situation is slightly different for the six-year-olds. They also show higher incidence rates at 8:00 a.m. and the noon hour. The highest incidence of discipline for the six-year-olds, however, comes at 4:00 p.m. Again, the rate of discipline for the period 4:00 p.m. to 7:59 p.m. is greater than at any other period during the day. The nine-year-olds do not show as many peak periods as do the younger children. There is a higher incidence of discipline around the noon hour and then a gradual rise in the rate of discipline from 4:00 p.m. through 8:00 p.m. For this group as well, the rate of discipline beginning at 4:00 p.m. and continuing until the end of

the day is greater than the rate of any previous period during the day. When the distributions are combined three periods emerge with the highest rate of discipline. These are the 8:00 a.m. hour, the noon hour, and the afternoon hours from 4:00 to 8:00 p.m. These hours correspond fairly well to the hours during which the children are preparing to go to school, returning for lunch and returning to school, and finally, are home from school for the rest of the day.

b. Days of week. The data were also examined for possible differences occurring in discipline on various days of the week. These data are presented in Table 1 (*Doc.*).

It is quite interesting to observe that although the children are not in school week-ends, and presumably spend more time at home then; Saturdays and Sundays, for the three-year-old group, do not demonstrate a significant increase in the number of incidents recorded. For the six-year-old group Thursday has significantly lower amounts of discipline than Wednesday, Friday, Saturday, and Sunday. For the nine-year-old group Wednesday, Thursday, and Friday exhibit significantly lower amounts of discipline than Monday, Tuesday, Saturday, and Sunday.

If most of the fathers spend more time at home on week-ends, no information about this is available, increased father-child contacts might increase the number of incidents. While some indications exist that the rate of discipline tends to rise on the week-ends the trend, however, is not clear cut. Thus a dichotomy in terms of week days versus week-ends is not particularly useful in explaining the differences obtained nor is there sufficient evidence to say that increased father-child contacts increases the number of disciplinary incidents.

c. Rate of recording. In order to estimate the faithfulness with which parents reported incidents, a mean rate of reporting per day by week of participation in the study was obtained. This information is given in Table 2 (*Doc.*).

When age groups are considered, the three-year-old mean difference between the first and second week is significant at the .05 level, while the difference between the first and third week is significant at the .01 level. For the six-year-old group the picture is approximately the same; the mean difference between the first and second weeks of the study is significant at the .05 level, while the mean difference between the first and third weeks is significant at the .001 level. In both of these age groups the difference between means for the second and third weeks was not significant. None of the differences between means for the nine-year-old group were significant.

For the total sample differences between the mean number of incidents per day for the first and second weeks, and for the first and third weeks, are significant at the .001 level. No significant differences are obtained in the mean number of incidents per day for the second and third weeks of the study.

The results would seem to indicate that the parents started the study with an initially high level of reporting which significantly (for the three- and six-year-old groups) dropped off at the end of the first week. There is a continuing decline, but this is not significant. Two hypotheses may possibly account for the decline of reporting over the three-week period. First, participation in the study may have had the effect of modifying the disciplinary behavior of the parents. Goodenough (3) demonstrated that parents in her study improved their control as a result of participation. While no direct evidence is available in this study to enable one to assess the effect of participation, it would seem likely that a similar phenomenon could occur with discipline as well. Second, although the parents were well motivated to carry on with the study, keeping records each day is a chore. It is interesting to note that in the group with the least number of incidents to report, no significant differences in the rate of reporting emerged. It is also possible to assume the initial interest may have resulted in a high rate of recording during the first week of the study which then leveled off as familiarity with the instrument increased and as the novelty of the situation for the parents decreased.

2. *Extrinsic Factors Associated with Discipline*

Extrinsic factors are those which are considered to be external to the disciplinary situation itself but which may play a rôle in the occurrence of discipline. These are general external conditions affecting the parent or the child separately and independently that lead to discipline in situations that ordinarily would not require such discipline.

a. *Sleep.* Loss of sleep may increase tension and sensitize the child to stimuli that would not otherwise affect him. Foster, Goodenough, and Anderson (2) have pointed out that lack of sufficient amounts of sleep has been associated with various emotional and physical maladjustments. They also conclude that such an explanation is warranted in a large number of instances. If the quantity and quality of sleep are related to sensitivity, then one might expect a relation to exist between these and the frequency with which children are disciplined. Among the factors that might affect the frequency of discipline could be the amount of sleep the child had the night before and whether the sleep the child did get was sound or restless.

(1). *Restless sleep.* The total number of days restless sleep occurred for all age groups and the percentage of that total on which discipline followed is given in Table 3 (*Doc.*). Accompanying these data are data relevant to the significance of the differences obtained between the age groups. To assess significance, the proportions of days on which discipline did and did not occur were compared for two age groups in a two by two contingency table. Three comparisons were made; between three-year-olds and six-year-olds, between six-year-olds and nine-year-olds, and between three-year-olds and nine-year-olds. Chi squares were obtained for each contingency table for 1 *df*, following McNemar (6, p. 203).

It can be seen that the occurrence of restless sleep tends to decline with age, but even at the three-year-old level it does not occur frequently. Despite this, when restless sleep does occur with the three-year-old, more often than not it will be followed by discipline. With the six- and nine-year-olds, however, the occurrence of restless sleep tends not to be followed by discipline. This reversal of tendency is highly significant when three-year-olds and nine-year-olds are compared. No other age comparisons were significant. There were no significant differences obtained in comparisons between sex or *SES* groups.

(2). *Amount of sleep.* Does the child experience more discipline the following day when he receives less than his usual amount of sleep. Table 1

TABLE 1
THE FREQUENCY AND MEAN NUMBER OF DISCIPLINARY INCIDENTS ACCORDING TO
WHETHER THE CHILD WAS ABOVE OR BELOW HIS MEDIAN OF SLEEP

Age	Below median sleep				Above median sleep				Critical ratio
	f	N	Mean	SD	f	N	Mean	SD	
3	422	40	10.55	6.62	368	40	9.20	6.52	.92
6	210	40	5.25	3.81	207	40	5.18	4.64	.07
9	104	40	2.60	2.76	119	40	2.98	3.64	.53

presents the frequency of disciplinary incidents, the number of families, and the mean number of disciplinary incidents per family according to whether the child was above or below his median sleep. Critical ratios are also presented for the difference between mean frequencies of discipline for above and below median sleep.

The data strongly indicate that very little relationship exists between the amount of sleep the previous evening and the frequency of discipline the following day. None of the differences obtained are significant, although for the three- and six-year-old groups there is a slight tendency for these children to have less sleep the evening preceding discipline.

b. Events. The possibility exists that various events occurring during the day may serve to heighten sensitivity to the point where both parent and child may become involved in disciplinary situations that otherwise might not arise. For example, having visitors in the home may increase the excitement level, carrying over into other situations after the guests have departed. Fatigue, after shopping or driving, may sensitize parents and children to situations that might otherwise have been ignored.

Space was provided on the Daily Information Sheet for the parent to check the occurrence of various events during the day. Table 4 (*Doc.*) presents the total number of days each of the events listed were checked and the proportion of the total on which discipline occurred.

For the three-year-old group, following the occurrence of an event, more days of discipline than of non-discipline occur. The situation is reversed for the six- and nine-year-olds with more days on which discipline did not occur following the checking of an event. To assess the difference between the age groups, for each event listed two by two contingency tables for each age comparison were constructed. These consisted of the proportion of days on which discipline did or did not occur. Following McNemar (6, p. 203), chi squares were obtained, and these are given in Table 4 (*Doc.*).

If one considers the number of days on which discipline did or did not occur, no one event can be associated with the occurrence of discipline. The same phenomenon occurs here as with the sleep behavior of the children. For the three-year-olds more days of discipline occur when the items are checked, and a reversal takes place for the six- and nine-year-old groups with more days on which discipline did not occur following the checking of items. While no sex or socioeconomic status differences are significant, in general there seems to be a significant relationship between the age of the child and the occurrence of discipline associated with various events. Even within the three-year-old group, if television is taken as an example, a large proportion (40 per cent) of the days television was watched was not followed by discipline. One would hesitate then to make any prediction as to whether or not discipline will occur if an event is checked for that day.

c. Health. Health is also a factor to be considered in the estimation of how frequently a child is disciplined. One assumption would be that the child who is ill may make more demands upon the mother, and therefore the rate of disciplinary incidents would increase. A different assumption which could be held is that the child who is ill is tolerated more and that behavior that ordinarily would be open to question by the parent would be ignored. Table 5 (*Doc.*) presents the total number of days different states

of health occurred and the proportion of those days on which discipline occurred.

Proportionately for the three- and six-year-olds, more days of discipline tend to occur when the child is ill. The proportion of days on which discipline occurs remains relatively constant for the nine-year-olds. It would seem then, that the younger the child, the greater is the likelihood of disciplining occurring when he is in ill health. Although health may be a factor in the occurrence of discipline, it is difficult to predict that discipline will occur when one knows the health status of a child on a particular day.

d. Hours under observation. It might be postulated that the frequency of reported incidents tends to increase with increased observation of the child. Accordingly, correlations were run between the mean hours observed per child over the three-week period and the number of disciplinary incidents over the same period. For the three-year-old group $r = .20$, for the six-year-old group $r = .11$, and $r = .08$ for the nine-year-old group. None of the correlations proved to be significant, indicating a lack of relationship between the mean hours observed and the frequency of discipline.

3. *The Content of Disciplinary Incidents*

Following the description of the length of disciplinary incidents intrinsic factors operating in the disciplinary situation will be described. Information was obtained about these by asking three questions. What was the child doing at the time the necessity for discipline arose? What difficulty precipitated the onset of the disciplinary action by the parent? What, according to parental report, was the issue involved?

a. Length of incidents. The distribution of disciplinary incidents and the mean time in minutes per disciplinary incident was obtained for each age group. The mean differences in duration of disciplinary incidents were also compared for the three age groups involved. These data are given in Table 6 (*Doc.*).

It can be seen that the average time per incident was significantly shorter for three-year-olds than for six-year-olds. While the time involved, on the average, in the disciplinary incident for the three-year-old is also shorter than that for the nine-year-old, this difference is not significant.

It can also be seen that the disciplinary incident, on the average, tends to be relatively short for all age groups. There is a concentration of disciplinary incidents within a 20-minute period. Three-year-olds have 92 per cent, six-year-olds 87 per cent, and nine-year-olds 90 per cent of the incidents falling within that time period. The modal incident for all groups takes

approximately 1-5 minutes in duration. In fact, slightly more than 60 per cent of all incidents for all groups are terminated within a five-minute period, and more than 85 per cent of all incidents are terminated within a 20-minute period. Thus discipline is over within a relatively short period of time. It would appear that the daily flow of the child's behavior is broken by relatively short, intermittent disruptions of social interactions. These episodes, although brief, can vary in intensity depending on the reasons for discipline and the actual or perceived intensity of the administration of discipline.

b. Child doing at the time. Statements of the mothers, with regard to what the child was doing at the time the disciplinary incident started, were categorized and coded. Originally 65 distinct behavior categories were found. Upon tabulation, however, it was found that this discrete a classification system was not feasible. Behavior categories were therefore combined in the following manner:

Play—solitary, siblings, parents, peers, with objects, general boisterousness, noisy, and not specified.

Eating—breakfast, lunch, dinner, snack, picnic, waiting at table, and eating—not specified.

Sleep—getting undressed for bed, general preparation and getting into bed, lying in bed awake, in bedroom—not otherwise specified, sleeping—wakes up, getting up, and nap.

Dressing—general morning preparations, getting dressed and/or undressed, changing clothes, washing and bathing, toileting, and grooming.

Television—

Activities—preparation for an activity, telephone, church, Sunday School, car, movies, shopping, visiting and/or being visited, spectator events including libraries and meetings, chores, practicing, reading, writing, singing, and other cultural activities including school work.

To and fro—going to the table, leaving the table, preparing to go to school, going to school, returning from school, preparing to leave home, leaving and entering the house.

Adult Interaction—medication, conversation, arguing, child making request of adult, adult making request of child, adult calling to child, and child doing something with or to any adult.

Miscellaneous—watching, waiting, doing nothing in particular, engaged in forbidden activity, searching or looking for an object, and in bed ill.

Not stated—

Two age groups at a time were compared in a two by two contingency table for each category listed. Chi squares were obtained to determine the significance of the differences between age groups. Table 7 (*Doc.*) gives

the distribution of families by age group using each of the categories. It can be seen that most families, regardless of age, report that the child was *playing* immediately prior to the onset of discipline. *Eating, sleep, and activities* were frequently indicated by all age groups as being the concern of children immediately prior to discipline.

For all categories there are no significant differences between sex or *SES* groups within age levels. When each of these categories is considered independently, there is a significant decrease with age in the number of children involved in *play* at the beginning of discipline. No significant differences emerge between sex and socioeconomic status groups within different age levels. Significantly more three-year-olds were *eating* at the time than were six- or nine-year-olds. The difference between six- and nine-year-olds is not significant.

Sleep behavior involved significantly more three-year-old children than nine-year-old children. The differences between three- and six-year-olds and between six- and nine-year-olds are not significant. The trend, however, seems to be in the direction of fewer involvements in this type of activity by older children.

The areas of *dressing, television, and activities* do not involve any significant differences between age groups. Significantly more three-year-old children were involved in the behavior category *to and fro* than were nine-year-old children. The differences between other age groups are not significant. For *adult interaction*, significantly more three-year-olds were involved than were nine-year-old children and significantly more six-year-olds were involved than nine-year-old children.

We have seen that significant differences arise between age groups in some of the categories examined. These differences are in the direction of a decrease with age. Essentially they seem to be a reflection of the fact that three-year-olds are disciplined more frequently than six-year-olds, and six-year-olds more than nine-year-olds. With more frequent discipline, families of the three-year-olds have more opportunity to be included within a particular category; thus the differences that do arise may be an artifact associated with the decrease in the frequency of discipline with age. The data do point out, however, that the children were concerned with daily routines and activities when the necessity for discipline arose.

c. *Difficulty that arose.* Parents reported a variety of reasons when they were asked to state what difficulty arose. Originally 89 categories were discriminated. It was found that this was too fine a breakdown, and the following categories were used in the final analysis:

Sibling relationships—quarreling, teasing, fighting and hitting (aggression), sharing possessions, threatening, and general interference with and/or disturbance of siblings.

Eating—refusal to eat, poor appetite, making a mess and/or playing with food, noise, playing, jumping up and down from table, use of utensils, manners, eating wrong foods or demanding foods, asking for food and refusing to eat it, and leaving table during the meal.

Sleep—delaying bedtime, in and out of bed, noise, refusal to go to bed or back to bed, refusing to get up, taking inappropriate objects to bed, remaining awake, playing instead of sleeping, and naps.

Dressing—selecting inappropriate clothing, changing clothes, getting clothing soiled or wet, proper care of clothing, grooming, self-help with clothing, refusal to wear clothing selected, refusal to get dressed or undressed or to put clothing back on, getting oneself dirty or wet, bathing, washing, teeth, and toileting.

Activities—difficulty with television, movies, church, Sunday School, practicing, difficulty in car and inappropriate behavior in car, and chores.

School—difficulty with school work in school, preparing to go to school, coming home late from school, refusal to go to school, unauthorized entering school, and taking material to and from school.

Health—medication—drugs, medication—appliances, and general health protection.

Inappropriate behavior—general restlessness and irritability, boisterousness and roughhousing, destructiveness, lying and stealing, rudeness, swearing and obscenity, tantrums-frustrations-crying, dangerous activities and forbidden behavior, overexcitement-fatigue-ill-humor, fighting and aggression, child afraid, jealous behavior, and nervous habits (thumbsucking, nose-picking).

Adult interaction—adult request refused by child, interference with adult activity, child request refused by adult, not getting adult permission or following adult requests or demands, persistent child demand of adult, refusal by child to compromise or cooperate, disobedience and non-compliance, adult possessions, and hitting (aggression) adult.

Home—neatness of home, coming home late, dawdling, accident, refusal to come in or not coming when called, refusal to put things away, and proper use of equipment.

Social—inappropriate behavior in social situation, unsocial behavior, showing off, teasing animals, social difficulties, interference with activities of others, insistence upon own rights, child ignored or passed over, difficulties with peers—sharing-aggression, and baby sitting.

Table 2 presents the distribution of families using the various categories subsumed under the difficulty that arose. Chi squares were obtained between age groups and between sex and socioeconomic groups within the age levels

for the various categories indicated. With the exception of the *home* category no significant sex differences are obtained, while over all categories no *SES* differences are significant.

TABLE 2
DIFFICULTIES PRECIPITATING DISCIPLINE BY FAMILY AND THE AGE OF THE CHILD

Difficulty	Age groups			Age comparisons (Chi square)		
	3	6	9	3-6	6-9	3-9
Siblings	31	24	21	2.85	.46	5.49*
Eating	33	18	10	12.17*	.01	26.60***
Sleep	31	22	12	4.52*	5.12*	18.15***
Dressing	30	24	14	2.05	5.01*	12.93***
Activities	9	15	17	2.14	2.08	3.65
Inappropriate behavior	28	25	14	.50	6.05*	9.82**
Adult interaction	35	29	18	1.95	6.24*	31.56***
Home	29	17	10	7.36**	.01	18.06***
Social	22	13	3	4.11*	6.33*	18.85***

* Significant at .05.

** Significant at .01.

*** Significant at .001.

When differences between groups indicating *siblings* as the difficulty preceding the disciplinary incident are examined, there is a significant difference between the three-year-old and nine-year-old groups, with the three-year-olds having more difficulty in sibling relationships. No other age differences are significant.

With the category *eating*, significant differences emerge between the three-year-olds and six-year-olds and between three-year-olds and nine-year-olds. There are no significant differences between the six-year-olds and nine-year-olds. Parents of three-year-olds seem to be attempting to train their children in proper eating habits and table manners. This difficulty tends to decrease with age and seems to become somewhat more stabilized by the age of nine although 25 per cent of the families of nine-year-olds still indicate incidents in this area.

The *sleep* behavior of children differs significantly from age to age in terms of being a precipitating factor in discipline. A possible explanation is that the routines associated with sleep have not yet become integrated or stabilized at the lower age levels. This process of integration with age seems to be independent of sex or the socioeconomic groups represented.

Dressing and the behavior subsumed under this category involves significantly more three-year-olds than nine-year-olds, and significantly more six-year-olds than nine-year-olds. The difference between three-year-olds

and six-year-olds is not significant. This data would further add to the impression that discipline arises from routines in which children are involved. When these routines become stabilized, the frequency with which they appear associated with discipline tends to decrease. This process is associated with age.

There is an increase in *activities* as a precipitating factor with age. The differences, however, between age groups are not significant.

Both the *school* and *health* categories had too few cases within cells to compute Chi squares reliably. Therefore, no analysis was made within these categories.

When *inappropriate behavior* is considered, significantly more three-year-olds display this behavior than do nine-year-olds and significantly more six-year-olds than nine-year-olds behave inappropriately. The difference between three- and six-year-olds is not significant.

There is a tendency for *adult interaction* to precipitate less discipline with age. The difference between three-year-olds and nine-year-olds is significant as is the difference between six- and nine-year-olds. The difference between three- and six-year-olds is not significant.

When *home* is considered, a consistent decrease in this factor occurs with age. Significantly more three-year-olds than six-year-olds, and significantly more three- than nine-year-olds are involved in behavior found in this category. The difference between six- and nine-year-olds is not significant. Significantly more three-year-old boys than girls ($X^2 = 4.51, p = .05$) are disciplined because of this factor. Other sex differences are not significant.

With regard to *social* difficulties, differences between all age groups are significant. More three- than six-year-olds, more three- than nine-year-olds, and more six- than nine-year-olds are involved in *social* difficulties. There is a consistent and significant decrease in this type of difficulty with age.

When one compares what the child was doing at the time the necessity for discipline arose and the difficulty that arose, there is a rather close relationship. For example, more three-year-olds just prior to discipline were concerned with eating behavior and proportionately more three-year-olds are described as having difficulty in this area. On the other hand, more nine-year-olds are shown as being engaged, just prior to discipline, in various activities, and they show proportionately more difficulty in this area than do the younger age groups. The difficulty that arises then, seems to spring from the immediately preceding activity of the child, and these activities are concerned with the ordinary processes of everyday living.

e. *Issue involved.* The parents were asked to indicate with a brief state-

ment the issues that were involved in the disciplinary incident. For many parents the issue involved was identical to the difficulty that arose, making it possible to use the same scoring system for the *issue involved* as was used for the *difficulty that arose*. The findings are similar to those concerning the difficulty that arose. Almost all the areas of daily living, eating, adult interaction, and other routines are the sources from which discipline springs.

4. *The Methods of Discipline*

a. *Methods of control.* In order to facilitate the record keeping required of the parents in the study, a mimeographed list of the different measures of controls and their exact definition was left with each parent. The parent was free to indicate on the *Observational Record: Discipline* sheet either by number or description the method or methods used in controlling the disciplinary situation. It was later found necessary to add four methods not originally given to the parents on the mimeographed list. The following is a complete listing of the categories used:

1. Appeal to emotions (make mother feel bad).
2. Appeal to humor.
3. Appeal to self-esteem.
4. Bribe.
5. Coax.
6. Deprive of food.
7. Divert child's attention.
8. Frighten or scare.
9. Ignore.
10. Isolate in a separate room or closet.
11. Praise.
12. Put to bed.
13. Reason.
14. Remove child forcibly.
15. Remove source of difficulty.
16. Ridicule.
17. Scold.
18. Shame—tell the child he is like a little baby.
19. Social approval.
20. Social disapproval.
21. Soothe.
22. Spank or slap (corporal punishment).
23. Take away privileges.
24. Tell the child that he will have to make up for what he did (restitution).
25. Tell the child that you won't love him anymore.
26. Threaten.

27. Other (parents were asked to specify—this category is now reserved for methods left unspecified).
28. Manual persuasion.
29. Order.
30. Substitution.
31. Shout.

Items 28 through 31 were added to the original list.

b. Methods per incident. The relationship between the frequency of discipline and the number of methods used within the space of a single disciplinary incident is given in Table 3. This table also presents the mean number of controls used per incident by age group and an examination of the significance of the difference between these obtained means.

TABLE 3
THE FREQUENCY AND MEAN NUMBER OF METHODS OF CONTROL USED PER INCIDENT
BY AGE GROUP

Number used	Age groups			Age comparisons—mean incidents per family (Critical ratio)		
	Three	Six	Nine	3-6	6-9	3-9
1 method	218	128	105			
2 methods	301	167	65			
3 methods	204	86	49			
4 methods	83	38	11			
5 methods	27	8	2			
6 methods	11	1	0			
7 methods	4	2	0			
Total no. of incid.	848	430	232			
Total no. of methods used	1993	932	436			
Mean no. of methods used per incident	2.35	2.17	1.88	3.00**	4.14***	5.88***
Standard deviation	1.33	.85	.96			

** Significant at .01.

*** Significant at .001.

Parents of three- and six-year-olds may use as many as seven different methods of control within the space of a single disciplinary incident, while parents of nine-year-olds may use as many as five within the space of an incident. The modal number of methods of control used is two for the three- and six-year-olds, while the modal number of methods used is one for the nine-year-olds. Thus we see that parents of three- and six-year-olds are more variable in their disciplinary behavior within a single incident than are parents of nine-year-olds.

When the mean number of controls per incident is considered, a significant decrease with age occurs. The idea that, on the average, parents tend to use

multiple rather than single controls is upheld. Further, this tendency decreases with the advancing age of the child, although at the nine-year-old level the use of multiple controls is still in evidence. One explanation that may be advanced is that parents of three-year-olds are continually exploring which method will do best, while at the nine-year-old level parents may know a little better which method they can utilize more effectively and which method the nine-year-old responds to best. Nine-year-olds not only have fewer disciplinary incidents reported, but within those incidents fewer methods of control are required. This decrease with age in the number of controls used per incident is also reflected in the percentage of the total number of controls used. The nine-year-old group accounts for 13 per cent of all controls reported in the study, the six-year-old group 28 per cent, and the three-year-old group is responsible for over half—59 per cent—of all controls used.

c. Role of mother. Does the mother tend to restrict herself to a rather narrow range of methods of control or is her behavior variable? Are there methods of control which are generally considered to "work better" than others? Data relevant to these questions are given in Table 8 (*Doc.*). This table compares the methods of control used by mothers of the three age groups used in this study according to the number of mothers using the various controls and by frequency of usage.

It becomes apparent at once that a large number of mothers report the use of *reasoning* and, in terms of rank, this remains in the first position regardless of the age of the child involved. For all groups *reasoning* retains the same rank when it is used second in a sequence of controls within an incident. As more controls are used per incident, *reasoning* tends to drop in ranking.

Again, for all age groups, *scolding* is in the second rank, although in terms of frequency it is used half as much as *reasoning* for all age groups. *Scolding* also maintains the second rank when it is used second in a sequence of controls within an incident. There is a tendency for it to increase as a third method used by mothers of three- and six-year-olds. Perhaps it should be pointed out that there may be difficulties distinguishing between reasoning and scolding; that reasoning with a child can easily turn into scolding. This is not to imply that the parents were not conscientious in their reporting, but that the discrimination necessary might vary in interpretation from parent to parent.

An interesting comparison can be made between age groups when one selects those controls which have been used by 50 per cent or more of the

sample, in this case, $N = 20$ for each age group. For the three-year-old group 11 controls, the first 10 controls listed plus *bribe*, are used by at least 20 families. For the six-year-old group three controls, *reasoning*, *scolding*, and *spanking*, are used by that many families, while for the nine-year-old group, *reasoning* alone is used by at least 50 per cent of the mothers. It would seem that the behavior of mothers of three-year-olds is much more variable than the behavior of mothers of older children.

The above comparison can be carried further. When the number of controls used by 25 to 49 per cent of the mothers ($N = 10-19$) is examined, nine, eight, and four controls are used respectively by mothers of three-, six-, and nine-year-olds. The data again indicates that variability in the disciplinary behavior of mothers tends to decrease with the age of the child involved.

All age groups are unanimous in reporting that *frightening* the child as a technique is not used anywhere in the sequence of controls. Furthermore, only one mother (that of a nine-year-old) reported the use of *withdrawal of love* as a technique, and this was preceded by two other methods.

It is interesting to note that while *reasoning* is ranked highly, it is not used as a technique by all the parents. The other techniques, not mentioned in this discussion are used by less than 25 per cent of the mothers.

How well do the mothers of the various aged children agree with one another? When groups three and six are compared, $\rho = .84$ ($p = .01$), for mothers of six- and nine-year-olds $\rho = .75$ ($p = .01$), while $\rho = .57$ ($p = .01$) for the three- and nine-year-old comparison. Mothers agree with each other at greater than chance expectancy. The question arises, do these correlations differ significantly from one another. In assessing the differences between groups, correlations were partialled out, since the same age group is involved in any two comparisons (6). When this is done, the difference between the three-six and six-nine correlations is not significant. The difference, however, between the three-six and the three-nine correlations is significant at .001, while the difference between the three-nine and the six-nine correlations is significant at .05. These results indicate that while mothers agree with each other at a level significantly greater than chance, there is a considerable difference in the patterning of responses. As the age differences between children increase, the indications are that the pattern of the disciplinary behavior of mothers differ to a significant extent.

d. Role of father. What does the father do in the disciplinary situation? Is his behavior highly variable or does he tend to restrict his disciplinary behavior in some manner? Data relevant to these questions is given in Table

9 (*Doc.*). This table compares the methods of control used by fathers of the three age groups used in this study according to the number of fathers using the various controls and the frequency of usage.

It becomes apparent at once that regardless of the age of the child fathers participate in the disciplinary situation to a lesser degree than mothers. Also apparent is the use of *reasoning* and *scolding* by fathers. As with mothers, these techniques are the ones most frequently used. The one change occurs in the nine-year-old group where *scolding* rather than *reasoning* is in the first rank.

In the three-year-old group alone, only one item, *reasoning*, is used by more than 50 per cent of the fathers. When one looks at the items used by 25 to 49 per cent of the fathers ($N = 10-19$), it is found that in the three-year-old group five items, *scolding*, *removing the child forcibly*, *spanking*, *coaxing*, and *threatening* are used, while in the six-year-old group three items, *reasoning*, *scolding*, and *isolation* are used. Any one method of control, in the nine-year-old group was used by less than 25 per cent of the fathers in the sample.

How well do fathers agree with one another? When fathers of three- and six-year-olds are compared, $\rho = .80$ ($p = .01$), for the six-nine comparison $\rho = .63$ ($p = .01$), while $\rho = .51$ ($p = .01$) for fathers of three- and nine-year-olds. Fathers agree with each other at a level significantly greater than chance. In order to determine whether these correlations differ significantly from one another, correlations were partialled out, since the same age group is involved in any two correlations (6). When this is done the correlations between three-six and six-nine, as well as the correlations between three-nine and six-nine do not differ significantly. The difference in correlations, however, between three-six and three-nine is significant at .01. The results indicate that the greater the difference between the ages of the children being punished, the more likelihood there is of the pattern of punishment of the fathers varying. One difficulty emerges in attempting to compare the rankings of controls used by fathers. Since so few fathers, relatively speaking, take part in the disciplinary situation, there are a great number of ties in the ranks, making the ranking procedure, at best, an extremely rough measure.

How well do mothers and fathers agree with each other? The rank order correlations between mothers and fathers for the three age groups range between .80 and .88. They agree with one another at a level significantly greater than chance. When the correlations are compared, no significant differences emerge. Generally speaking, although the father participates to

a lesser degree than does the mother in discipline, when he does participate he tends to use the same methods.

c. *Role of both parents.* Does the disciplinary picture change when both parents are operating together? Table 10 (*Doc.*) presents data comparing the methods of control used by both parents jointly of the three age groups used in this study, according to the number of parents using the various controls and the frequency of usage.

The situation does not change. *Reasoning* remains the method most frequently used, although in the nine-year-old group it is tied in ranking with *scolding*. *Scolding* maintains the second rank with parents of six-year-olds, dropping to the third rank with parents of three-year-olds. For the three-year-old group, *coaxing* is in the second rank.

In the three-year-old group alone, only one item, *reasoning*, is used by slightly more than 50 per cent of both parents operating jointly. When one looks at the items used by 25 to 49 per cent of both parents ($N = 10-19$), it is found that in the three-year-old group three items, *coaxing*, *ignoring*, and *threatening*, are used, while in the six-year-old group one item, *reasoning*, is used. Again, for the nine-year-old group no one item is used by this many parents. Any single method of control for the nine-year-old group is used by less than 25 per cent of the parents.

It should be pointed out that both parents operating at the same time do not frequently discipline the child. As a matter of fact, they operate jointly about as frequently as the father does alone. Mothers discipline the children approximately 74 per cent of the time, fathers 12 per cent, and both operating together 12 per cent. Others enter into the disciplinary picture about 2 per cent of the time. It is, therefore, clear that the bulk of discipline taking place within the family is administered by mothers, and almost never by anyone other than the parents.

How well do parents of the different aged children in this study agree with each other? The rank order correlation between the parents of three- and six-year-olds is .66 ($p = .01$), for the six-nine comparison $\rho = .60$ ($p = .01$), while $\rho = .37$ ($p = .05$) for parents of three- and nine-year-olds. Parents disciplining together agree with each other at a level significantly greater than chance expectancy. However, these correlations are of a lower magnitude than those obtained for mothers or fathers independently. In order to determine whether these correlations differ significantly, correlations were partialled out, since the same age group is involved in any two correlations (6). The difference between the three-six and six-nine, as well as the difference between the three-nine and six-nine correlations is not

significant. The difference in correlations, however, between three-six and three-nine is significant at .05. The results again indicate that the greater the difference between the ages of the children being punished, the more probable the pattern of punishments of the parents will vary. Again a difficulty emerges in attempting to compare the rankings of parents operating jointly. Since so few parents, relatively speaking, operate jointly in the disciplinary situation, a great number of ties appear in the ranks, making the ranking procedure an extremely rough measure.

f. Role of others. Persons other than the mother or father are occasionally put in the position of disciplining the child. These are usually close relatives such as grandparents, siblings, aunts, and uncles, although occasional household help (maids, domestic help, and baby sitters) are also involved. As may be expected, these persons are not frequently involved in the disciplinary situation. For this group there is approximately a 2 per cent participation. It should be pointed out, that there is more participation by others in discipline with the younger child, and this participation decreases with increase in the age of the child.

Since this group does not participate to any large degree no data will be reported for it. It is of some interest, however, to see what methods these persons use when they are involved in discipline. *Reasoning* is used most frequently with the three- and six-year-old groups, but it does not occur in the nine-year-old group. *Scolding* is second in rank for the three- and six-year-old groups, while in the nine-year-old group it is tied with *spanking* for first position.

Again, behavior is more variable on the part of those disciplining three-year-olds. Fourteen different methods were used with the three-year-olds, 13 with the six-year-olds, and five with the nine-year-olds. The three-year-olds then not only are faced with more discipline when compared to the older age groups, but are also faced with a greater number of different controls even when discipline is administered by others.

g. Ten most frequent methods used. The frequency distributions of mothers, fathers, both operating jointly, and others were combined and the controls in the first 10 ranks are given for each age group separately in Table 4. It can be seen that *reasoning* and *scolding* maintain the first two ranks for all age groups. *Coaxing* also tends to remain fairly high for all groups. *Spanking*, on the other hand, is high for the three-year-old group, drops down in rank for the six-year-old group, and disappears from the first 10 methods for the nine-year-old group.

Threatening remains fairly high as a technique for all age groups, as

does the tendency on the part of the adult to ignore the situation or the child. Adults attempting to use diversion rank this technique higher with three-year-olds than with six-year-olds, and it does not appear among the first 10 techniques utilized by the adults dealing with nine-year-olds. *Taking away a privilege* from a child is not used too frequently with the three-year-old group, appears in the lower ranks of the six-year-old group, and achieves the third ranking in the nine-year-old group. *Humor* appears at the low end

TABLE 4
THE FIRST TEN METHODS OF CONTROL BY RANK AND AGE OF THE CHILD

3-year-old group		6-year-old group		9-year-old group	
Method	Rank	Method	Rank	Method	Rank
Reason	1	Reason	1	Reason	1
Scold	2	Scold	2	Scold	2
Coax	3	Coax	3	Take away privilege	3
Spank	4	Threaten	4	Coax	4
Divert	5	Ignore	5	Self-esteem	5.5
Threaten	6	Isolate	6	Threaten	5.5
Ignore	7	Spank	7	Ignore	7
Remove child forcibly	8	Divert	8	Remove difficulty	8
Isolate	9	Take away privilege	9		
Remove difficulty	10	Humor	10.5	Humor	10
		Order	10.5	Isolate	10
				Social disapproval	10

of the first 10 methods for both six- and nine-year-old groups and is not used even this frequently as a technique with the three-year-olds. *Removing the child forcibly*, however, is a technique which appears in the top 10 techniques for the three-year-old group only, while appeal to the *self-esteem* of the child is found in the first 10 ranks of methods for the nine-year-old group only.

Six of the first 10 techniques used by parents are common to all groups, *reasoning*, *scolding*, *coaxing*, *threatening*, *isolating*, and *ignoring*. *Spanking* and *diverting* appear among the first 10 controls for the three- and six-year-old groups. One technique is common to the three- and nine-year-old groups, that of *removing the difficulty*. Two of the techniques used, *taking away a privilege* and appeal to *humor*, appear among the first 10 methods of control for the six- and nine-year-old groups.

On the basis of the data presented then, it appears that differences between age groups in terms of the methods of control used by parents in the disciplinary situation are small. It seems that a relatively small number of techniques tend to recur in almost all disciplinary situations regardless of the age of the child. These first 10 techniques account for more than 70 per cent of the total number of methods used.

5. *The Outcome of Discipline*

a. Resolving the issue. For each disciplinary incident that occurred, the parent was asked to state the outcome by checking one of the following: (a) Did the child give in? (b) Did the child get his own way? (c) Was the issue compromised? (d) Was the issue ignored or glossed over? (e) Was the issue left unsettled? (f) Was it impossible to assess what happened to the issue? (g) Was it impossible to assess because the control was used after the incident was closed?

The frequency and the mean number of disciplinary incidents per outcome is given in Table 11 (*Doc.*). In addition, comparisons between age groups per category are also given.

It can be clearly seen that the child yields in overwhelming numbers. When age groups are compared, differences between the mean number of incidents in which the child yields are highly significant. Yielding on the part of the child decreases with age to a significant extent. This may be a function of the number of incidents reported for each age group rather than a function of the age of the child, since in all groups this tendency is high. In 67 per cent of the three-year-old incidents, 71 per cent of six-year-old incidents, and 59 per cent of the nine-year-old incidents does the child yield. These data are not quite consistent with the decline in number of incidents with age that the significance figures show.

In relatively few of the incidents (one per cent) is the *issue ignored* by the adult. *Compromising the issue* or *child gets own way* does occur, but in a relatively small number of cases (approximately nine and eight per cent respectively). Also in a relatively small number of cases is discipline applied after the incident was closed. While the differences between the three-year-old group and older age groups are significant, applying discipline after the incident is closed infrequently occurs with older children. If one, on the basis of the frequency with which the child yields the issue in the disciplinary situation, can equate yielding with success in the disciplinary situation, then these parents have been quite successful. Success, however, cannot be measured this easily. The child may yield in the situation only because of the superior force and/or position of the adult and it may not give information as to whether the child will adjust his behavior in some future similar situation.

b. Behavior of the child. Responses to the question, "How did the children act after the controls were used?" were tabulated and categorized. The following is the categorization scheme used:

1. *Normal.* All responses that indicated the child "was just fine," resumed activity, was his "usual self," or "settled down."
2. *Continuance of behavior.* All responses that indicated no reaction, or that the child was unimpressed or "casual."
3. *Active cooperation by child.* Responses that indicated the child was cooperative, understood the issue, accepted substitutions, accepted the situation or has his attention diverted.
4. *Compliant behavior.* Responses that indicate the child apologizes, is resigned, sorry or contrite.
5. *Pleasantly toned behavior.* Responses that indicate the child is pleasant, laughed, was pleased or happy.
6. *Subdued and unhappy behavior.* Responses that indicate the child is subdued, quiet, thoughtful, unhappy, morose, tense or sorry for himself.
7. *Angry and irritable behavior.* Responses which indicated the child was angry, fumed, irritable, cross, mad, sullen, sulked, pouted, acted annoyed and left the scene.
8. *Aggressive and resistant behavior.* Responses that indicate aggression, resistance, argumentation, belligerence, and defiance.
9. *Crying and temper tantrums.*
10. *Not stated.* Responses that are omitted or impossible to determine from the record.

Table 12 (*Doc.*) presents the frequency and the mean number of disciplinary incidents associated with various categories.

The most frequent response of the three-year-olds is *crying*, and this is significantly greater than either the six- or nine-year-old, while the most frequent response for both the six- and nine-year-olds is *anger*, although the differences between groups are not significant. If discipline is perceived as a frustrating situation for the child, crying and anger responses can be expected. However, relatively few parents indicate that the children are compliant, subdued, or aggressive as a result of the discipline administered. More of these types of behavior might be expected if the disciplinary situations were always perceived as frustrating. The number of incidents in which the behavior of the three-year-old is labeled pleasant is rather surprising; one might expect more anger or aggressiveness. Either the child has learned to inhibit overt responses such as aggression to avoid further punishment, or the disciplinary situation cannot always be equated with frustration. The possibility exists that when the parent makes the child change his behavior, the change itself may lead to greater satisfaction, either because of the reward inherent in the situation or because of the parent's pleasurable response after the change has been made.

Generally speaking, the behavior of children after discipline has been

administered is quite variable. This, taken in conjunction with the data that indicates that the child yields quite frequently in the disciplinary situation, can possibly be interpreted as indicating that yielding by the child may be the result of successful application of methods of control in the disciplinary situation by the parents. Another interpretation lies in the nature of the parent-child relationship. Since the child is dependent upon the parent, and since the relationship established between parent and child has tremendous import, the child yields in the situation, inhibiting aggressive responses on his part, to avoid counter aggression and to maintain the relationship.

c. Duration of child's behavior. The consideration of duration becomes complicated since many of the parents did not state the time in terms of numbers, but indicated the passage of time in many different ways. Those incidents which were time referent were used to obtain the mean duration of the behavior per incident. The mean duration for three-year-olds is 10.84 minutes, for six-year-olds 15.02 minutes, and for nine-year-olds 13.34 minutes. The mean duration of the behavior after the control was used is significantly shorter for three-year-olds than for six-year-olds. Other differences obtained are not significant. The modal duration of the after-behavior for all groups is one to five minutes. Along with the means obtained for the three age groups, indications are that the duration of after-behavior is relatively short for all age groups. Further, over 50 per cent of all disciplinary incidents have the after-behavior terminated within five minutes. It should be remembered that the behavior being timed here includes *normal* and *pleasantly toned* behavior as well as some of the more negative aspects of the reactions to discipline such as *anger* or *aggression*.

While on the average the duration of the after-behavior is brief, there is a large amount of variability. The range as well can be considered extensive, from less than one minute to more than one hour. Those incidents which are not time referent would also appear, from the descriptions alone, to be more extensive than the mean times indicate. On the basis of the data obtained, it would be difficult to predict what the length of a particular reaction to discipline will be.

6. Adjustment and Parental Attitudes

a. Adjustment and discipline. One measure that lends itself to obvious comparison is the relationship between adjustment and the frequency with which children are disciplined. Correlations were therefore computed between the scores individual children obtained on the adjustment scales used (Personality Profile and the *H-O-W* Behavior Rating Schedules) and the

total number of disciplinary incidents recorded for them while they were under observation. This was done for each of the scales used. Correlations were run between the scores assigned to children by the teachers as well as by the parents on the scales mentioned. These data are given in Table 13 (*Doc.*) by age and sex, and in Table 14 (*Doc.*) by age and socioeconomic status.

Although some of the correlations in Tables 13 (*Doc.*) and 14 (*Doc.*) are significant, no age, sex, or socioeconomic trends become apparent. One hesitates, therefore, to postulate the existence of a relationship between adjustment, as measured, and the frequency with which children are disciplined in the home.

If one were to assume that the adjustment score reflects the frequency with which discipline will be encountered by the child, then, on the basis of the above data the relationship is higher for parents than teachers. Those correlations of the teachers that reached a criterion of significance were in the "wrong" direction; that is, when the teachers rated the children as being well adjusted, there were more disciplinary incidents occurring in the home. Parents, in general, were in the "right" direction, although few of the correlations reached significance. Why should the relationship be higher for parents than teachers? Perhaps the answer lies in the fact that the rating of adjustment given to the child by the parent includes the knowledge of how frequently their children get into difficulty. This knowledge may be taken into account, implicitly, by the parent when she makes her rating of adjustment. That discipline is not the only factor involved in rating the child is shown by the general low level of correlations obtained.

The parents were able to reflect the frequency of discipline successfully for the six-year-olds on the Personality Profile ($p = .05$), for the nine-year-olds on the *H-O-W* III ($p = .01$) and for the nine-year-olds on the *H-O-W* IV ($p = .05$). The correlation for the three-year-old group rated by the teachers is significant ($p = .05$). This correlation, however, is in the "wrong" direction. Whenever the teacher rated the child as being well adjusted, there were significantly more disciplinary incidents in the home.

Do parents and teachers differ significantly in their correlations between the frequency of discipline and adjustment ratings? The data were further examined for significant differences between the correlations obtained for parents and teachers. Few of them turned out to be significant. On the Personality Profile, one significant difference emerged; the difference in coefficients for the six-year-old girls is significant at the .02 level. On the *H-O-W* III, differences obtained for the nine-year-old girls are significant at

the .02 level and the differences for the total nine-year-old group is significant beyond the .001 level. On the *H-O-W* IV, the difference between teachers and parents for the nine-year-old *SES* I group is significant at the .02 level, while the difference for the total nine-year-old group is significant at the .05 level.

One might conclude from the data presented that the relationship between the frequency of discipline and adjustment is higher for parents. It should be pointed out, however, that the parental measures of adjustment are not necessarily independent measures since such factors as discipline may affect her rating of the adjustment of her child. If the teacher's rating can be considered as an independent criterion, since there may not be any relationship between the frequency of discipline in school and at home, then one can conclude that there is essentially no relationship between good adjustment and the frequency of discipline; that the frequency with which a child is disciplined is independent of his adjustment rating.

b. Minnesota Scale of Parent Opinions. The *Minnesota Scale of Parents' Opinions* was filled out by each mother at the beginning of the study. This instrument consists of two subscales, one relating to home practices, the other traits the mother would desire in children. Each subscale consists of 40 items, each item having a weighted response. A score is obtained for each individual by adding the weights assigned to each response. The scale is so constructed that high scores indicate opinions held that agree with "good" concepts of child development. A minimum score of 80 points, would indicate a parent who reacted strongly in the "wrong" direction, while a maximum score of 400 points would indicate a parent who reacted extremely in the "right" direction.

The parents in this study avoided the extreme positions, yet were consistently within the bounds of opinion which is in keeping with current thought about child development practices. The mean scores tended to decrease with the increase in age of the child. Parents of three-year-olds had a mean of 310.75, of six-year-olds a mean of 304.52, and parents of the nine-year-olds a mean of 298.45. These differences are not significant.

The similarity between groups is emphasized when one finds the mean ratings assigned to items by the parents. Here the ratings are consistently toward the higher ("more correct") end, yet, avoid extreme positions. The mean rating per item for the parents of three-year-olds is 3.88, for the six-year-olds 3.81, and for the nine-year-olds 3.73. These differences are not significant. The age trend is, of course, apparent in these scores.

Within the range of scores obtained on the *Minnesota Scale of Parents'*

Opinions, interest is centered upon the relationship that might exist between opinion, as measured by this instrument, and the frequency of discipline encountered in the home. Accordingly, correlation coefficients were obtained between the scores on the Minnesota Scale and the number of incidents reported per family.

Correlations between total group scores and the frequency of discipline are not significant; they are all low but positive indicating the possibility of some relationship. Two of the correlations obtained are significant; the nine-year-old boys— $r = .55$, $p = .05$, and the nine-year-old SES II group— $r = .63$, $p = .01$. Very little relationship seems therefore to be present between parental attitude and the frequency of discipline, subject, of course, to the limitation of the small number of subjects used.

D. SUMMARY AND CONCLUSIONS

The purposes of this study were (a) to describe the disciplinary process as it takes place in the home, (b) to examine relationships between extrinsic factors and discipline, and conditions of social interaction and discipline and, (c) to determine the effects of discipline on the child.

The relatively homogeneous sample consisted of 120 mothers of children, 60 boys and 60 girls equally divided among three age levels, three, six, and nine years. Representation was equal from Groups I and II of the Minnesota Scale for Paternal Occupations.

Data obtained from mothers consisted of the following: (a) Information about the family. (b) An estimate of how frequently various methods of disciplinary control were used. (c) A measure of parental attitudes toward child-rearing practices. (d) Two adjustment ratings of the child. (e) A daily record, for twenty-one consecutive days, giving information pertinent to health, activity level, habits, and events occurring in the life of the child. (f) A descriptive account of each incident of discipline occurring during the twenty-one days on forms specially prepared to equate observations from parent to parent. Data obtained from teachers consisted of two adjustment ratings for each child, identical in form to the ones made by the mothers.

1. *General Findings*

1. On all of the dimensions examined, the primary factor associated with discipline was the age of the child. Occasional differences were found between sex and socioeconomic groups, but these occurred rarely. There was a significant decrease with age in the frequency of discipline. One can conclude then, that the age of the child is the greatest single predictor of discipline.

2. Discipline arises in those situations where the child is concerned with routines of daily living, establishing sibling and adult relationships, and in displaying behavior that adults deem inappropriate. In these, the probability exists that the difficulties arising will decrease with age.

3. Mothers are the persons who are most frequently responsible for disciplining the child. While fathers do participate, they do so to a minimal extent. There is a great degree of similarity in the disciplinary controls that all mothers and all fathers used, regardless of the age of the child involved. There are also some indications that parents shift the methods of disciplinary control they use with the age of the child.

4. While a large number of different disciplinary controls are used, general preference, in terms of use, is expressed for the use of reasoning as a technique. There are also some indications that a shift takes place with the age of the child from manual or physical techniques to those which are more verbal in nature.

5. When discipline does take place it is likely to be relatively short in duration. When considered with the average rate of discipline per day per child, discipline can be seen as a short, sporadic interruption of the daily flow of behavior.

6. As the day progresses there is greater likelihood for discipline to occur. This was indicated by the number of incidents occurring as one peak period succeeded another during the day.

7. While there are some indications that discipline tends to increase on weekends, there are also other days during the week when discipline tends to increase as well, indicating little or no relationship between discipline and the day of the week.

8. The behavior of children in reaction to the application of disciplinary controls seems to be varied. This behavior ranges from pleasant, normal behavior to crying and aggressive behavior. The younger child tends to be more variable in his reactions to the application of discipline than does the older child.

9. A low positive relationship exists between the frequency of discipline and adjustment ratings of the child. This relationship, however, is not significant. This indicates a trend in the right direction, but for this sample, however, there is very little relationship between the frequency of discipline and adjustment.

10. It was found that a low but positive relationship exists between the degree to which parents held opinions concerning child-rearing practices and

the frequency with which they disciplined their children. This relationship, however, was not significant.

2. Discussion

a. Development and the decline of discipline. What possible factors can account for the decline in frequency of discipline with age? Several factors associated with growth and development may offer explanations. As the child grows older, along with growth in other areas, there is an increase in communication skills. Not only is the child better able to make his needs and desires known, but he also becomes increasingly aware of attitudes and desires of others. Since discipline decreases as a function of age, it is possible that the decrease in the frequency of discipline and growth in communication are related. Being able to make one's needs and desires known may enable parent and child to interact with one another without discipline. Although making needs and desires known may lead to more friction between parent and child, since the statement of the need may occasionally come into conflict with the parent's desires, the probability is that increased communication leads to smoother interaction between parent and child.

The effects of increased communication may also be seen in parental behavior. Parents can also make themselves better understood. They may not have to resort to action, words can guide and direct the behavior of the child. This also might lead to discipline, since the statement by the parent may be in conflict with the child's desires, but again the probability is that increased communication leads to smoother interaction between parent and child.

The probability exists that the increase in communication serves to make the child understand and be aware of parental attitudes. With increased awareness and understanding of parental attitudes and feelings, the behavior of the child can be modified, since the awareness of existing attitudes and feelings may serve to inhibit or facilitate behavior before the necessity for discipline arises. Thus the child need not come into conflict to be disciplined, but his awareness of the attitudes themselves takes on a disciplinary function.

How does the child become aware of attitudes that then direct him in his behavior? The effects of learning must be taken into account. Developmentally, we can say that the child proceeds from rudimentary concepts to those which are more differentiated. These, of course, are not only consequences of learning and communication, but of growth in intelligence as well. The child is better able to see, feel, and understand situations from a more intellectually integrated frame of reference. This growth enables the child

to make increasing use of cues and partial cues, and makes him more perceptive of parental attitudes and actions. In addition, with growth and exposure to a large variety of learning situations the child establishes his own set of behavioral standards which have been partially established on the basis of parental standards. Since guiding one's own behavior is one of the prime purposes of discipline, then the degree to which the child has absorbed standards with which to guide his own behavior becomes a measure of the success of the discipline used. The evidence of this study points indirectly to the fact that this absorption or internalization of values has taken place. Not only does the frequency of discipline decline with age, but there is also a decline with age in almost all of the specific issues involved which precipitate discipline. The older child is better able to get along with others, better able to follow routines, and better able to participate in activities without their leading to discipline.

Another factor which also partially accounts for the decline in frequency of discipline with age is that parental evaluation of the child may change as the child grows older. The parent expects different behavior with the older child. There is a different evaluation of the child's abilities and capacities, both in terms of his behavior and his increasing ability to use self-control. From a parental point of view, the older child should be able to guide himself in many areas and he is given the responsibility for doing so. Because of changes in the evaluation of the child, the rôles parents see themselves as playing in the disciplinary situation may change by taking into account the present status of the child.

As the child grows older, parents have a more adequate basis for comparing the child's present status with a picture they might have of him as an adult. The parents have trained the child according to what they believe is proper. It is possible that as the child learns he becomes more like the type of person the parents want him to become. Where the perceptions of present and future status of the child are not too dissimilar, that is if the child is developing in the expected direction, the need or desire to modify the child's behavior may be lessened, resulting in a decrease of discipline.

b. The necessity for discipline. Why and how does discipline arise? The data of this study point conclusively to the fact that extrinsic factors have very little bearing on the occurrence of discipline. The mere fact that a child watches television, for example, or takes longer than usual to go to sleep the night before, does not seem to make him more prone to disciplinary action. The data do show, however, that the younger child is more sensitive to these extrinsic factors. Although watching television *per se*, for

example, will not enable one to predict with any degree of accuracy that discipline will follow, with the three-year-old there is a greater likelihood that discipline will follow watching television than for older children. Thus disruptions of routines and various occurrences tend to cause a greater amount of disequilibrium in the life of the younger child, although the disequilibrium in and of itself may not be sufficient to cause discipline. Possibly this may be associated with the fact that the younger child still has to grow in many areas, and with growth he becomes less sensitive to these distractions or is less affected by them.

The stated causes of discipline do not revolve about extrinsic factors in the environment. No mother, for example, reported that discipline occurred because the child did not get enough sleep. Rather, the causes centered about problems of sibling relationships, adult-child interactions, inappropriate behavior of the child, eating, sleeping, dressing, activities centered about the home, and social difficulties. All of these problems involve interactions between people. The difficulties that arise are disruptions of interpersonal relationships. Thus the child is in active interaction with his world, and as a result of this interaction he becomes involved in situations leading to conflict with the parent. It is what the child does to others and what others do to him in social interaction that results in conflicts, the need for resolution of conflicts, and the use of discipline. With growth and development the child learns to adjust his behavior to the varied demands of his environment, and the result is a decrease in the number of conflicts with parents. Not only is there a general overall decline in the number of conflicts, but almost all of the specific areas studied demonstrate this same age relationship to a significant extent.

c. Disciplinary techniques. The number of different disciplinary techniques used by the parents decreased with the age of the child. Not only was the three-year-old disciplined more frequently than the six-year-old, and the six-year-old disciplined more frequently than the nine-year-old, but the number of disciplinary controls used with these age groups decreased in similar order. Why the older child faced less discipline and fewer methods of disciplinary control may be explained by several factors. The growing capacity of the child for language usage and communication may influence the parent in eliminating and/or relying less heavily upon a large number of techniques, particularly those which are non-verbal in nature. When the parent can communicate verbally or vocally, there may be no need to resort to other disciplinary measures. As the child grows older he is building up a repertory of behaviors that are more appropriate to situations in which he finds him-

self. His behavior is changing through learning and it becomes more predictable by the parent and/or the child is more successful in avoiding many situations in which a greater number of disciplinary controls would be used. The child also learns to predict the behavior of the parent and can anticipate, and thus prevent, the use of disciplinary controls. This may be related to the greater use of cues and partial cues by the older child.

Another trend in the data that is of interest is that with increasing age there was greater resort to techniques which are more verbal in nature. The parent of the three-year-old may, for example, resolve the conflict by simply picking up the child and removing him from the scene. This method is forceful, non-verbal, and somewhat direct. The greater use of forceful techniques with the younger child again is an indication of the essential lack of communication between parents and children. As the child grows older the parent is better able to cope with a variety of situations in a less forceful manner. While the use of verbal techniques was evident in all the age groups studied, and reasoning was the technique most frequently used, there was a decrease in the use of forceful, non-verbal techniques with age.

While the mother is the most frequent disciplinarian, there is a great degree of similarity between mothers and fathers in their use of disciplinary techniques. There is also a great degree of similarity between all mothers, and a great degree of similarity between all fathers, regardless of the age of the child. Despite these similarities, however, there is some evidence which points to the fact that some shifting takes place in the use of disciplinary measures. This shifting is associated with the age of the child. That is, parents are attempting to pattern their practices to the ages of their children. This again indicates that parents may be aware of the growth and development taking place within their children and take this into account in the disciplinary situation. Thus parental behavior, as well as child behavior, is changing.

e. Effects on the child. Several questions need to be answered in discussing the effects of discipline on the child. First, what effect did discipline have upon the modification of behavior? In other words, was discipline successful? Second, what were the immediate emotional effects? Third, are any effects manifested in the adjustment of the child from a long-term point of view?

It has already been mentioned that the criterion of success of discipline is whether or not the child has been able to absorb a system of values by which he can come to guide his own behavior. It has been stated that there is some indirect evidence to uphold the belief that this is taking place. Another attempt to assess the success of discipline can take into account how

the issue was resolved. The data of this study point to the fact that the vast majority of children yielded the issue when discipline was applied. On the surface then, parents did accomplish their purposes since children did yield in the disciplinary situation. While the yielding of the issue is not the best criterion for assessing the success of discipline, when it is coupled with the fact that there was a demonstrable decrease with age in specific issues causing discipline, with decreases in frequency of discipline with age, with decreases in the number of disciplinary methods used with age, and shifts in the disciplinary patterns of parents with the age of the child, yielding the issue can serve as one index in a manifold of criteria for judging the success of the disciplinary process.

When behavioral reactions to the administration of discipline are examined, two types of variability emerge. First, over all age groups, there is variability in the types of responses that all children display. That is, the reactions to the application of discipline are varied rather than specific. Second, an age difference in variability of response is also in evidence. It was found that the behavior of the three-year-olds was more variable than the behavior of older children. They demonstrated significantly greater amounts of normal and pleasantly toned behavior as well as aggressive and crying behavior. This would indicate that although conflict is involved in the disciplinary situation, it cannot be considered a paradigm for frustration.

It is much more difficult to assess the long-term effects of discipline on the child. If one can make the assumption that the disciplinary patterns within a family are reflected in the adjustment ratings the child receives, then one method is available with which effects can be measured. The data demonstrate a low, positive, but insignificant relationship between adjustment and the frequency with which children are disciplined. This indicates a trend in the right direction, but for this sample, however, there is very little relationship between the frequency of discipline and adjustment. This latter finding holds true for all age groups and is independent of the person making the adjustment rating.

3. Suggestions for Future Research

One of the omissions of this study is the failure to take into account possible relationships between the affectional state of the mother at the time discipline was administered and the type and frequency of discipline. In addition, some assessment of the general personality patterns of the mothers would have proved valuable. It is possible that disciplinary patterns of mothers are related to immediate emotional feelings as well as being related

to their personality patterns. While most of the emphasis in studies on discipline have been upon the child, much more information is needed on the rôle of mothers and fathers in discipline, not so much concerned with what they do, but with how they feel.

This study assessed disciplinary practices as they arose in conflict situations. It has not answered the question, "How does discipline work in non-conflict situations?" Is it possible to isolate such factors as the tone of voice, gestures, facial expressions, and the like and assess their effect on the behavior of children? Possibly a combination of the observational technique used in this study, observations by trained observers in the home, and intensive interviewing of the parents based on the records obtained would expand our knowledge even further.

Questions concerning the rôle of communication and its effect on discipline need more extensive investigation. If the extent of communication between parents and children can be measured, then it might be posited that a relation exists between communication and the type and frequency of discipline.

Finally, much more is known about parental practices in the negative aspects of discipline. More emphasis needs to be placed on the use of rewards in the disciplinary situation.

REFERENCES

1. FLESH, R. F. *The Art of Readable Writing*. New York: Harper, 1949.
2. FOSTER, J. C., GOODENOUGH, F. L., & ANDERSON, J. E. The sleep of young children. *J. Genet. Psychol.*, 1928, **35**, 201-218.
3. GOODENOUGH, F. L. *Anger in Young Children*. Minneapolis: Univ. Minnesota Press, 1931.
4. INSTITUTE OF CHILD WELFARE. *The Minnesota Scale for Paternal Occupations*. Minneapolis: Univ. Minnesota, 1950.
5. JACKSON, P. W. Verbal solutions to parent-child problems, *Child Devel.*, 1956, **27**, 339-349.
6. MCNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949.

Child Research Council
University of Colorado School of Medicine
4200 East Ninth Avenue
Denver, Colorado

DEVELOPMENT OF A MATERNAL BEHAVIOR RESEARCH INSTRUMENT*

Laboratory of Psychology, National Institute of Mental Health

EARL S. SCHAEFER, RICHARD Q. BELL, AND NANCY BAYLEY

This research arose from a need for a set of constructs which would permit the organization and quantification of unstructured descriptive data on mother-child interaction. After a discussion of a theory of concept development and a brief review of previous concepts and methods used in organizing and quantifying maternal behavior, we will present the origin and orientation of the set of concepts utilized here. The method of rating will be described as well as the results of this attempt at conceptual refinement.

A. THEORY OF CONCEPT DEVELOPMENT

Lazarsfeld (13) begins his insightful discussion of the process of concept development with the statement: "All social sciences meet with the problem of making inferences from simple observations to more complex 'things.'" He states that concepts such as friendliness have been included by logicians (9) under the term of "dispositional concepts" which "are defined as concepts which do not refer to a directly observed characteristic but rather to a disposition of some objects to display specific reactions under specified circumstances." A set of observed items of behavior is used to specify the meaning of the more general dispositional concept. Kaplan (12) has also discussed the problem of specification of meaning of a scientific term. He concluded that the definition of a scientific concept consists of the specification of a set of indicators which are positively correlated with one another. He states that "the specification at any stage is a provisional one, both as to the indicators included, and the weights associated with them." One possible implication of this statement is that ordinary speech and the conceptual systems used by a culture are an implicit step toward the development of more exactly defined, communicable, quantifiable, scientific concepts. This would provide a rationale for organizing the data of mother-child interactions in accordance with hypothetically important everyday concepts which can be more exactly defined by specifying the behavioral content which they subsume. These discussions of concept development (12, 13) are consistent with an attempt to organize the multitudinous behaviors which are included in mother-child interaction into a limited number of general concepts. Each

* Received in the Editorial Office on April 24, 1958.

of these general concepts would be defined by a specified set of behaviors and also would be assumed to relate to a wider population of behaviors of which the specified behaviors would be a sample.

B. REVIEW OF SELECTED STUDIES

The Fels Parent Behavior Rating Scales which were developed by Champney (6) and Baldwin, *et al.* (3, 4) are similar to the rating scales developed in this study. The Fels scales consist of 30 rating scales on parent behavior and family variables which are to be used by a trained observer after a visit to the home. High inter-rater and intra-rater reliabilities have been reported (4). The scales have been intercorrelated (3), factor analyzed (14, 20), and have been correlated with developmental data on the children (2). Although good results have been obtained through use of these scales, extensive training of the home visitor is required, and it has not been demonstrated that the scales can be utilized to quantify interviews or other observations outside the home.

Another method of measuring mother-child interaction was developed by Merrill (17) for quantification of observations of mother-child interaction in a standard playroom situation. The mother's behavior was recorded in 32 categories and analyses were made of the following: lack of contact, structurizing, structurizing a change in activity, teaching, interactive play, helping, directing, interfering, criticizing, coöperation, and noncoöperation. The method has been described sufficiently so that it can be used by others.

Sears, *et al.* (23), in a study of 40 mother-child pairs have contributed another conceptualization and quantification of maternal behavior as revealed by interviews. Their concepts, which arose from hypotheses concerning the antecedents of dependent and aggressive behavior of children, included the following—scheduling; severity of weaning; eating frustration; toilet training frustration; bedtime nurturance; nurturances when mother is busy; father nurturance; nurturance when child is sick or injured or upset; severity of toilet training; sickness-danger frustration; orderliness-cleanliness frustration; punitiveness; responsiveness to aggression. They were successful in quantifying these variables and demonstrated relationships between maternal interview variables and the development of dependent and aggressive behavior in children. This approach has been extended in a study by Sears, Maccoby, and Levin (24) of 379 mothers who were each interviewed for approximately two hours. From these interviews 188 variables were coded or rated. The data were analyzed in many ways to determine antecedents and consequents of various maternal interview variables.

In a cross-cultural study of child training by Whiting and Child (27) five systems of behavior—oral, anal, sexual, dependence, and aggression—were independently quantified relative to initial indulgence, severity of socialization, and age of socialization. Although significant correlations were found between initial indulgence, severity of socialization, and age of socialization on each variable, few correlations were found between systems of behavior. Relationships between other cultural variables and their indices of child training were demonstrated.

Many other studies have been summarized by Symonds (26), Radke (19), and Orlansky (18), and others are included in a bibliography by Heinicke and Whiting (8). These studies encompass a wide variety of conceptual schemes that have been used in the study of mother-child interaction. Despite the great amount of previous work in this area, we were motivated to develop a new set of measures for both theoretical and methodological reasons.

C. MOTIVATION FOR THIS STUDY

The theoretical motivation for our work came from the ideas of theorists such as Adler (1), Sullivan (25), Horney (10), George Mead (16), and others who stress the context of interpersonal relationships in which personality develops. More directly, the conceptual scheme which we have used has been derived from the intuitive concepts of psychologists, psychiatrists, and social workers who have studied mother-child interactions in various clinical situations. Case reports from these settings often implicitly employ a conceptual scheme approximating the one which is formalized in the rating scales presented here. Of course, the concepts used are limited by our theoretical bias and degree of insight as well as by the data which were available for analysis. A collection of data which provides more information would permit addition of other relevant concepts.

The methodological motivation for our study was a wish to develop a set of rating scales defined in behavioral terms which would be communicable and would permit reliable ratings by relatively unskilled personnel. The method used, which is described below, gives an efficient procedure for measuring any concept that can be clearly defined by specifying relevant behaviors.

D. SAMPLE AND NATURE OF DATA

The material utilized in developing this scale comes from the Berkeley Growth Study which has been described by Jones and Bayley (11) and Bayley (5). (These earlier reports give information on the organization of the total research program in which the data were collected.) We quote

from (11): "The 61 cases originally included in this study were 'normal' Berkeley children born in hospitals between September 25, 1928, and May 15, 1929. The families selected included only white, English-speaking parents who could be regarded as permanent residents of Berkeley and who were willing to coöperate in bringing their children to the Institute for the required series of examinations. In consequence of the selective factors involved, the group tends to be somewhat above the average in measures of socioeconomic status (parental occupation, income, and education)." The original sample was somewhat reduced as a few people left the area and a very few refused further coöperation; to some extent the case loss was replaced by a few additional subjects.

The data on which the first form of the Maternal Behavior Research Instrument was developed were notes on observations of 56 mothers seen at the Institute of Child Welfare in testing sessions during the first three years of the child's life. Tests were scheduled once a month for the first 15 months and every three months thereafter until the child was three. An observer who was present during the testing session made written observations¹ of the interactions of the mother with the examiners, the child, and any other persons present. The amount of data available on each mother varied with her communicativeness, the amount of significant behavior as noted by the observer, and the number of regularly scheduled sessions which the mother attended. The majority of the mothers were present at 10 to 20 sessions. Several cases were not rated because the mothers had not attended regularly or had left the study before sufficient data for ratings were available.

The second set of data to be analyzed was composed of reports of interviews with 34 of the same mothers in the home situation, done by a clinical psychologist² during the summers of 1938, 1939, 1941, and 1942. Although two interviews were available on the majority of the cases, there were several cases for which only one interview was available. The interviews were structured to provide information on the following:

I. Ratings of: *a.* House exterior. *b.* Neighborhood. *c.* Yard. *d.* Living room.

II. A narrative descriptive picture of the home.

III. Ratings of parents: Use of language: *a.* Fluency. *b.* Vocabulary. *c.* Cultivation. *d.* Intelligence.

¹ Most of these were written by H. Elizabeth Neall, a graduate student in Psychology, additional notes were made by experimenters, L. V. Wolff, M.D., and N. Bayley.

² Dr. Pearl Bretnall Meissner.

IV. Narrative account of evaluation of mental ability of child by parent.

V. Narrative account of ability of parents to evaluate individual differences and age changes in intellectual functions.

VI. Narrative account of parents' attitude toward parent-child similarities.

VII. Narrative account of parental evaluation of inter-sibling differences.

VIII. Narrative account of parents' ambitions for children.

IX. Narrative account of parental personality characteristics.

X. Narrative account of the following sources of intellectual stimulation: *a.* Reading interests. *b.* Movies—type and frequency. *c.* Radio. *d.* Parental contact with school. *e.* Special opportunities—educational excursions, hobbies, training in skills. *f.* Opportunities for social contact—children, adults.

The focus of the interview was upon data that would pertain to intellectual growth of the child although personality variables of the parents were also included. Some interviews were with the mother alone, others with both mother and father present. The child was absent or present depending upon the home situation at the time of the interviewer's visit. The material recorded was detailed, graphic, and insightful, and preserved the individuality of the subjects.

E. PROCEDURE IN THE CONSTRUCTION OF THE RATING SCALES

The method employed in developing the rating scales, although it has been implicitly used by many investigators, was formally described by Furfey in 1926 (7) and was utilized by McDonough (15) in quantifying the behavior of children in the classroom situation. The procedures used in developing the scales are described below. By using these same procedures, one can construct additional ones, relevant to the material to be rated.

In developing a set of measures for a behavioral domain, one could proceed from specific behaviors to general concepts or vice versa. One could collect a large number of specific behaviors which are relevant to a domain, type the individual behaviors on cards and then sort the cards into groups which are judged to be logically and psychologically homogeneous. From the set of behaviors which are expected to be highly correlated with one another, a general concept could be derived. Such a procedure would not require an explicit, pre-conceived set of concepts, but only a choice of relevant behaviors from which concepts could be developed. The method for proceeding from the general concept to specific behaviors that define the concept would be as follows. If one were concerned with a concept such as suppression of sexu-

ality in the child, one could, from his experience, describe specific behaviors which were relevant to the concept and which might be observed in maternal behavior.

Most of the scales of this study were developed by choosing the major concepts first and then selecting trait-actions which defined the concepts. The trait-actions used to define the concepts of this study were written after reading the protocols to be rated. An effort was made to select for use trait-actions that could be rated from the available data. There are several advantages in specifying trait-actions or behaviors which define the general, more abstract concept. The raters are given a common definition; the concept as used can be more completely communicated to others; and higher reliability may be expected from combining the several ratings of the trait-actions into a total score for the concept.

The first step in rating the records of this study was to acquire a set of implicit norms for use in making the ratings. The rater familiarized himself with a representative sample of the data. This practice insured more direct comparability between raters, than would be expected if ratings had been made on the basis of previously acquired norms. With this background of the protocols in mind, the rater then re-read each individual record carefully immediately before rating it. The general concept was not rated, but each of the trait-actions was rated on a scale which ran from: Not at All True "1," through Average "4," to Extremely True "7." The raters were encouraged to attempt to obtain a normal distribution of scores for the group. Extensive training of raters was not necessary since the concepts and trait-actions of the scales utilize non-technical language and only require such judgments as are commonly made about interpersonal relationships. The raters were told to rate overt rather than inferred characteristics. In this way they avoided conflicts that might result from diverse hypotheses about psychodynamic origins of particular behaviors. A preference for rating of observed rather than inferred characteristics is dictated by the hypothesis that the covert reactions of the mother to the child will be revealed by some aspect of her relationship with the child and will therefore be included in the pattern of the overt behavior rating scales.

Each scale was analyzed for rater reliability and internal consistency of trait-actions. Since the judges were instructed to omit a rating if there was no available evidence on that trait-action, it was possible to examine the score sheets and to eliminate the consistently unrated trait-actions. Scatter plots were made of the ratings of each judge with every other judge to determine the extent of agreement. If the judges showed little agreement, the trait-action was omitted.

Another test of the usefulness of a trait-action in defining the general concept was to note whether ratings on a given trait-action were correlated with the other trait-actions that defined the concept. If the trait-actions were not correlated with one another, the investigators had not succeeded in defining a homogeneous concept with a set of observable trait-actions. It was occasionally found that several trait-actions were highly intercorrelated while others were essentially unrelated. Those trait-actions which did not correlate with the majority in a scale were eliminated from that scale.

After elimination of inefficient trait-actions, interjudge reliability was calculated by Pearson product-moment correlations. For each judge, a subject's total score on the concept was determined by adding the scores for the trait-actions that defined the scale. There was marked variation both in mean scores and in variability of scores given by different raters. Despite instructions to use this sample as a reference group when rating, the inter-rater differences in mean scores are probably due to differences in norms of the raters. The variability of the judges may be partially due to differences in willingness to give extreme scores and willingness to attempt a discrimination of a small difference between subjects. Since these scales attempted relative rather than absolute measurement, each judge's distribution of total scores on each scale was transformed into a standard score distribution which equated means and variances of the judges. After this transformation to standard scores it was possible to calculate inter-judge reliabilities and to combine the scores assigned by the judges into a total score on a scale for each individual.

F. RESULTS

Reliabilities of the Maternal Behavior Research Instrument: Form I. Quantification of Behavior Observations (see Appendix *A*) were computed by correlating the transformed scores of each judge with those of each other judge for each scale. For these correlations it was possible to estimate the reliabilities of scores obtained by summing the ratings assigned by all three raters. For several scales on which one rater did not give ratings that varied sufficiently for reliable measurement, that rater's scores were not included. Because we eliminated both trait-actions that did not vary with other trait-actions defining a scale and trait-actions with inadequate reliabilities, the reliabilities for Form I are maximal. Since this scale was developed from materials which were not intended for this type of organization, observations that were systematically gathered to permit ratings with this conceptual scheme might yield higher reliability. One additional qualification of the

estimated reliability of the combined scores should be noted. If a rater had omitted rating one subject, the average score of the other two judges was used as this subject's rating in calculating scores. As such omissions were

TABLE 1
INTER-JUDGE RELIABILITIES^a OF THE MATERNAL BEHAVIOR RESEARCH INSTRUMENT
FORM I: QUANTIFICATION OF BEHAVIOR OBSERVATIONS
(Decimal points omitted)
N = 56

Scale No.	Correlation between judges			Combined reliability ^b	Name of scales
A-B	A-C	B-C			
1	77	81	63	89	Ignoring
8	72	75	67	88	Financial Stress
16	71	69	75	88	Perceive Child as Burden
12	67	67	73	87	Punishment
23	74	70	62	87	Emotional Involvement
24	66	70	64	86	Use of Fear to Control Child
3	76	54	66	85	Cooperativeness
17	74	61	60	85	Fostering Dependency
30	71	55	70	85	Intelligence
20	72	52	66	84	Competitiveness
14	66	57	65	84	Equalitarianism
18	71	56	62	83	Anxiety
10	68	61	58	83	Rejection of the Homemaking Role
31	67	51	67	83	Expression of Affection
11	58	59	67	82	Irritability
9	70	49	62	82	Positive Evaluation
7	69	60	51	82	Sociability
4	67	57	53	81	Poor Physical Health of Mother
32	60	53	63	81	Achievement Demand
15	68	48	59	81	Concern about Health of Child
26	71	50	52	80	Autonomy of the Child
29	48	63	60	80	Over-conscientiousness
22	64	55	50	79	Mood Swings
28	50	50	64	78	Strictness
19	— ^c	64	—	78	Punitiveness
13	58	48	47	76	Intrusiveness
6	50	49	44	73	Narcissism
5	44	45	51	72	Dependency of Mother
2	—	56	—	72	Self-Abasement
21	44	50	36	70	Negative Emotional State
27	—	51	—	67	Excessive Contact
25	—	—	—	—	Suppression of Aggression

^a Scores were normalized before calculating inter-judge reliability.

^b Spearman-Brown Formula applied to average inter-judge correlation.

^c Insufficient data to estimate reliability.

rare, the estimates should be approximately accurate with the limitations described above (Table 1).

The reliabilities of the ratings of Form I, reported in Table 1, vary from .70 to .89 with a median reliability of .82. It should be emphasized that

these are inter-rater reliabilities and that the validity of the observation or the degree of relationship between two independent sets of observations on these subjects cannot be determined from the data available.

The Maternal Behavior Research Instrument: Form II. Quantification of Interview Data (see Appendix *B*), was revised from Form I in order to make the trait-actions age appropriate and to make the scale appropriate for the interview material which has been described. The original concepts of Form I and the trait-actions defining them were retained whenever possible. Trait-actions which were added to define the same concepts were hypothesized to measure the same dimensions as the trait-actions in Form I that defined the concept. Again the records were rated by three judges, only one of whom had read the observations used in rating Form I. An inspection of the results of the ratings on Form II revealed high homogeneity in both the reliability of the trait-actions and the extent to which they appeared to measure the concepts they were chosen to define. An explanation of this finding might be that we had profited from our experience in Form I and had, as a result, chosen more appropriate trait-actions. A decision was therefore made to retain all trait-actions that had been used to define the concepts. The judges' total scores on a concept were again transformed to a standard score distribution and the inter-judge reliabilities and also the combined reliabilities were again determined. Apart from the unreliable variable of Dependency of the Mother, the range of combined reliabilities, reported in Table 2, was from .75 to .95 with a median combined reliability of .85. It should be noted that no attempt was made to maximize these reliabilities by eliminating trait-actions. It seems highly probable that if interviews with mothers were structured properly to gain maximum information about the variables included, it would be possible to obtain higher reliabilities.

The reliabilities reported here are interrater reliabilities derived from ratings of interview notes from one interviewer. These data do not permit an estimate of the reliability which would be obtained by correlating independent ratings of notes of two independent interviews. However the relationship of the ratings of the 0-3 year observation notes with the independent ratings of the 9-14 interview notes for 31 subjects who were studied at both periods furnish evidence of validity of these judgments (21). Additional investigation of the correlations of these ratings of maternal behavior, too complex to include here (22), has also supported the construct validity of this conceptual scheme.

G. SUMMARY AND CONCLUSIONS

A brief discussion of the process of concept development and of a method for developing measures of concepts has been presented. Measures for a set of concepts which are relevant to an interpersonal theory of personality

TABLE 2
INTER-JUDGE RELIABILITIES^a OF THE MATERNAL BEHAVIOR RESEARCH INSTRUMENT
FORM II: QUANTIFICATION OF INTERVIEW DATA
(Decimal points omitted)
N = 34

Scale No.	Correlation between judges			Combined reliability ^b	Name of scales
	A-D	A-E	D-E		
11	83	88	87	95	Marital Happiness
16	77	84	76	92	Positive Emotional State
2	79	80	75	91	Cooperation
4	82	76	72	91	Sociability
6	79	76	73	90	Rejection of Homemaking Role
1	72	81	65	89 ²	Ignoring
7	78	67	73	89	Irritability
27	80	68	65	88	Positive Mother-Child Relationship
28	70	74	67	88	Wish to Control
17	80	67	57	86	Emotional Involvement
14	62	74	67	86	Anxiety
15	65	74	62	86	Punitiveness
18	76	63	56	85	Use of Fear to Control Child
24	73	68	54	85	Achievement Demand
5	58	72	59	84	Positive Evaluation
10	53	76	58	83	Equalitarianism
12	60	56	61	81	Perceives Child as Burden
20	62	50	63	81	Excessive Contact
25	46	64	62	80	Communicativeness
8	47	64	60	80	Social Isolation
22	67	52	49	79	Intelligence
23	68	49	44	78	Expression of Affection
19	50	66	43	77	Autonomy of the Child
21	60	60	38	77	Strictness
26	54	58	45	77	Withdrawal
9	40	66	51	77	Intrusiveness
13	37	56	44	72	Fostering Dependency
3	38	00	07	56 ^c	Dependency

^a Scores were normalized before calculating inter-judge reliability.

^b Spearman-Brown Formula applied to average inter-judge reliability.

^c Only Judges A and D were used.

development were developed. The concepts, many of which are commonly used in clinical studies in child guidance clinics, are exactly defined by specifying the behaviors which they include. The utility of this method of quantifying both observations and interviews was demonstrated by the relatively high interjudge reliabilities of ratings.

Evidence of the construct validity of these scales has been presented else-

where (21, 22). We present them as promising tools, with the hope that they will be tested further. The two sets of rating scales could be used as they are reported here. Since they were developed for a particular set of data, other investigators must judge whether they could be used successfully on other data in their present form. However, new scales can be developed, or the present scales can be adapted, with the method which has been described.

REFERENCES

1. ANSBACHER, H. L., & ANSBACHER, R. R., Eds. *The Individual Psychology of Alfred Adler*. New York: Basic Books, 1956.
2. BALDWIN, A. L. Socialization and the parent-child relationship. *Child Devel.*, 1948, **19**, 127-136.
3. BALDWIN, A. L., KALHORN, J., & BREESE, F. H. Patterns of parent behavior. *Psychol. Monog.*, 1945, **58**, Whole No. 268.
4. ———. The appraisal of parent behavior. *Psychol. Monog.*, 1949, **63**, Whole No. 299.
5. BAYLEY, N. Mental growth during the first three years: A developmental study of 61 children by repeated tests. *Genet. Psychol. Monog.*, 1933, **14**, 1-92.
6. CHAMPNEY, H. The measurement of parent behavior. *Child Devel.*, 1941, **12**, 131-166.
7. FURFEY, P. H. An improved rating scale technique. *J. Educ. Psychol.*, 1926, **17**, 45-48.
8. HEINICKE, C., & WHITING, B. B. *Bibliographies on Personality and Social Development of the Child: Pamphlet 10*. New York: Soc. Sci. Res. Council, 1953.
9. HEMPEL, C. *Concept Formation in the Empirical Sciences*. Chicago: Univ. Chicago, 1952.
10. HORNEY, K. *The Neurotic Personality of Our Time*. New York: Norton, 1937.
11. JONES, H. E., & BAYLEY, N. The Berkeley Growth Study. *Child Devel.*, 1941, **12**, 167-173.
12. KAPLAN, A. Definition and specification of meaning. In P. F. Lazarsfeld & M. Rosenberg, Eds. *The Language of Social Research*. Glencoe, Ill.: Free Press, 1955.
13. LAZARSFELD, P. H., Ed. *A conceptual introduction to latent structure analysis*. In *Mathematical Thinking in the Social Sciences*. Glencoe, Ill.: Free Press, 1954.
14. LORR, M., & JENKINS, R. L. Three factors in parent behavior. *J. Consult. Psychol.*, 1953, **17**, 306-308.
15. McDONOUGH, M. R. The empirical study of character: Part II. *Studies in Psychology and Psychiatry*. Washington: Catholic Univ. Press, 1929. (Vol. II, No. 4.)
16. MEAD, G. H. *Mind, Self, and Society*. Chicago: Univ. Chicago Press, 1934.
17. MERRILL, B. A measurement of mother-child interaction. *J. Abn. & Soc. Psychol.*, 1946, **41**, 37-49.
18. ORLANSKY, H. Infant care and personality. *Psychol. Bull.*, 1949, **46**, 1-48.
19. RADKE, M. J. Relation of parental authority to children's behavior and attitudes. *Univ. Minn. Instit. Child Welf. Monogr.*, 1946, No. 22.
20. ROLF, M. A factorial study of the Fels Parent Behavior Scales. *Child Devel.*, 1949, **20**, 29-45.

21. SCHAEFER, E. S., BELL, R. Q., & BAYLEY, N. Quantification of maternal behavior and consistency of mother-child interaction. *Amer. Psychol.*, 1946, 11, 404. (*Abst.*)
22. SCHAEFER, E. S. Organization of maternal behavior and attitudes within a two dimensional space—an application of Guttman's Radex theory. *Amer. Psychol.*, 1957, 12, 401. (*Abst.*)
23. SEARS, R. R., WHITNEY, I. W. M., NOWLIS, V., & SEARS, P. S. Some child rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monog.*, 1953, 47, 135-234.
24. SEARS, R. R., MACCOBY, E., & LEVIN, H. Patterns of Child Rearing. Evanston, Ill.: Row, Peterson, 1957.
25. SULLIVAN, H. S. Conceptions of Modern Psychiatry. Washington, D. C.: Norton, 1948.
26. SYMONDS, P. M. The Psychology of Parent-Child Relationships. New York: Appleton-Century, 1939.
27. WHITING, I. W. M., & CHILD, I. L. Child Training and Personality. New Haven: Yale Univ. Press, 1953.

Child Development Section

National Institute of Mental Health

Bethesda 14, Maryland

APPENDIX A

MATERNAL BEHAVIOR RESEARCH INSTRUMENT:
Form I. Quantification of Behavior Observations

-
1. Does this mother ignore or reject her child?
 1. Does she often comment on how much extra work or trouble the child is?
 2. Does she tend to "leave the situation" during the examination as though she is glad the baby is in someone else's hands?
 3. Would she be willing to have others assume most of the responsibility for care of the child?
 4. Does the mother seem to know very little about the child?
 5. Does she tend to overlook the needs of the child?
 6. Does she give the impression that the child is not necessarily her principal interest?
 7. Does she fail to show much beyond polite interest in the child during the examination?
 2. Does this mother seem self-abasing?
 1. Does she make unnecessary excuses for how her child is dressed?
 2. Does she excessively justify missed appointments, lateness for appointments, or other failures to meet obligations incurred as a part of the study?
 3. Does she ask to be reassured that her child-rearing practices are adequate?
 4. Does the mother feel inadequate to cope with the child's behavior problems or physical care?
 5. Does the mother express feelings of personal inadequacy?
 6. Are there signs the mother can't handle the child because she is afraid to say "no" to it?
 3. Does this mother seem cooperative overtly? (Exclude officiousness, interventions, self-abasing cooperation.)
 1. Does she help examiners with physical and mental tests in a way that indicates a desire to help primarily?
 2. Does this mother seem to share the examining situation?

3. Is this mother sensitive and considerate of others?
4. Does she perceive the problems of the examiners in handling her baby?
5. Is her handling of husband or relatives indicative of sensitivity to their problems?
6. Does she accept the goals of the study understandingly?
7. Does she avoid talking or socializing when it might be disturbing in the examination?
4. Is this mother physically healthy?
 1. Is she in need of medical attention?
 2. Does she fatigue easily?
 3. Does she seem lacking in endurance and resistance to disease?
5. Is she a dependent mother?
 1. Does the mother seek support and reassurance from the experimenter?
 2. Does she demand a great deal of help from servants, relatives and friends?
 3. Does she seem more concerned about satisfying her own needs than those of the baby?
 4. Does she seem unable to handle even minor crises or illnesses alone?
 5. Does she frequently ask advice about child-rearing or family problems?
 6. Does she ask the examiners to make decisions for her?
 7. Does she mention friends, husband, in-laws or relatives in a way which indicates she must have people to rely on?
6. Does she try to draw attention to herself?
 1. Does she bring up the matter of her own education, social position or other matters in a way indicating a desire to impress others?
 2. Does she talk about or show off the baby primarily as a way of drawing attention to herself in a positive fashion?
 3. Does she cooperate with the study mainly because it makes her and her child the object of special attention?
7. Does this mother tend to be a social person?
 1. Does there seem to be more verbal communication with this mother than with others?
 2. Does she converse about the husband and family events in an easy, sociable way?
 3. Does the mother converse freely about her own reaction to events and her opinions, attitudes, and feelings?
 4. Does she try to establish friendly social relations with the examiners?
 5. Does she have a variety of skills in social contacts which indicate effort directed toward the establishment of friendly relationships with others?
 6. Does she tend to elicit positive personal and social reactions from the examiners?
8. Does this mother seem to be subject to financial strain?
 1. Does she talk about the difficulty her husband is having with his business?
 2. Is she worried about whether she can get things for her children that are usual for most children?
 3. Is she worried about providing food and clothes for her children?
 4. Is it necessary for her to work so as to alleviate the financial strain?
 5. Does her husband have the kind of job which fails to provide any real basis of security for the family?
9. Does the mother tend to have a positive evaluation of the child as a person?
 1. Does the mother think the child is attractive in appearance by her standards?
 2. Does the mother think the child is intelligent in ways she likes?
 3. Does the mother think the child has a desirable emotional disposition?
 4. Does the mother typically approve of the child's behavior?

5. Does the mother think of the child as sociable with adults and other children in ways she herself values?
6. Is there an absence of fault finding in the mother's behavior relative to the child?
10. Does this mother seem to reject the role of homemaker?
 1. Does she complain frequently about inadequate domestic help or the need for it where it does not exist?
 2. Does she frequently complain of being tired in connection with house-keeping activities?
 3. Does she use social engagements primarily as a way of escaping from her homemaking role?
 4. Does she speak of trips, visits, and other activities outside the home as though she sees them as means of getting a respite from homemaking?
 5. Does she ever use phrases which indicate a feeling that she feels held down or shut-up in the home?
11. Does this mother tend to be irritable?
 1. Does she complain that her child gets on her nerves?
 2. Does she comment on her own lack of patience?
 3. Does she talk about her frustration and irritation with the child's behavior?
 4. Does she seem to have little tolerance for the child's behavior in the testing situation?
 5. Is she abrupt and rough in her handling of the child or is her voice harsh and unsympathetic?
 6. Is she irritable because of the child's behavior or emotional reactions during the tests?
12. Does the mother believe in punishment as an effective method of influencing the child's behavior?
 1. Does the mother punish the child before he is able to talk?
 2. Does the mother punish the child for failures in toilet training?
 3. Does the mother punish the child because of his eating or sleeping habits?
 4. Does the mother spank the child when he is negativistic?
 5. Does the mother spank the child when he cries or is emotionally upset?
 6. Does the mother spank the child in order to get him to cooperate with her own or the examiner's instructions?
 7. Is the mother's punishment severe?
13. Does the mother seem to be unaware of the fact that a child has a mind of his own and that he should be able to do his own thinking without forceful intrusion by the mother?
 1. Does the mother seem to be training the child to think of himself as the mother's possession?
 2. Does the mother seem to think that she should be able to direct what he will think or say at any time?
 3. Does the mother's approval of the child seem to be strictly dependent upon the degree to which he adapts his thinking and actions to her wishes?
 4. Does the mother seem to think that she should know what the child is thinking or doing at all times?
14. Does the mother tend to relate to the child as an equal?
 1. Does the mother play games with the child in an informal way?
 2. Does the mother attempt to talk with the child at his level?
 3. Does the mother tend to reduce emphasis upon age and role differences between mother and child?
 4. Does the mother tend to avoid a directive, commanding role?
 5. Does the mother enjoy spending time with the child?

15. Does this mother seem to be preoccupied with her child's health?
 1. Is she quite fearful of the child catching a cold?
 2. Does she show real concern over minor illnesses or insignificant defects (bowleggedness, etc.)?
 3. Is she concerned greatly about the child's growth?
 4. Is she concerned greatly about the child's weight?
 5. Is she concerned greatly about the child's diet?
16. Does the mother perceive the child as more a burden and inconvenience than a source of happiness?
 1. Does the mother complain of great difficulty in caring for the child?
 2. Does the mother talk about the tension and anxiety of child-rearing?
 3. Does the mother complain of the fatigue of housework and caring for children?
 4. Does the mother talk about interference of the child with other activities and pleasures?
 5. Does the mother feel caring for the child is a burden rather than a joy?
 6. Does the mother show few signs of enjoying the child's activities?
 7. Does the mother talk primarily about problems she has in caring for the child?
17. Does this mother tend to baby her child or foster dependency in him?
 1. Does she often do things for the child he could do for himself?
 2. Does the mother wish to help the child with difficult problems in the test situation?
 3. Is she afraid he might be hurt if he plays with other children?
 4. Does the mother seem reluctant to have the child perform certain tasks or be measured because it will be upsetting to him?
 5. Does the mother see the child as weak, helpless, and as needing excessive help, care and support?
 6. Does the mother tend to anticipate the child's needs excessively and shelter it from even normal effort?
18. Does this mother appear overtly anxious during the session?
 1. Does she seem upset when the child cries?
 2. Does she seem upset when the child refuses to perform?
 3. Does she seem disturbed about the child's capacities on the mental and/or physical tests?
 4. Does she seem afraid that the child is being hurt?
 5. Does she seem tense or unduly constrained during the session?
19. Does the mother seem punitive and unkind?
 1. Does she take an "it serves you right" attitude toward any of the child's accidents or upsets?
 2. Does she talk in a negative vein about friends, husband, or associates?
 3. Does she show tendencies toward open attack on the baby when in conflict?
 4. Does she describe the child critically in such a way as to indicate she is gaining satisfaction from the criticism?
20. Does the mother show competitive, domineering behavior in the family and examining situation?
 1. Does she try to influence the scoring of the test in a way which indicates strong achievement drives?
 2. Does she try to rationalize failures so that the baby will look the way she wants it to look?
 3. Does she see the entire examining situation as one which she has to control in some way?
 4. Does she make efforts to change the social situation of the examination so that she is on the same level with the examiners?

5. Are there indications that she tries to get the baby to do things her way even in little matters in which the baby might just as well have its own way?
 6. Does the mother mention information on her own achievements in a way to indicate competitive behavior?
21. Does she tend toward negative emotional states?
1. Does she typically seem gloomy?
 2. Does she typically seem detached and inwardly absorbed?
 3. Does she typically seem dull and lacking in emotionality?
 4. Does she seem to be unresponsive at times?
 5. Does she occasionally fail to show adequate awareness of what is going on around her?
22. Does she show mood swings?
1. Does she show changes in cheerfulness or gloominess within or between sessions?
 2. Does she show signs of emotional lability in her home situation?
 3. Does she show changes in tension and relaxation within or between sessions?
 4. Does she vary between warmth and reserve in her discussion of how she and her child are getting along?
 5. Does she show changes in patience and frustration tolerance within or between sessions?
23. Extent and intensity of emotional and behavioral involvement with the child.
1. Is the mother very much interested and involved in the child's test behavior?
 2. Does the mother have a tendency to play, talk, and generally interact with the child during the test?
 3. Is the mother eager to quiet and comfort the crying child?
 4. Does the mother have frequent and close physical contact with the child?
 5. Does the mother spend a great amount of her time with the child?
24. Does the mother attempt to use fear as a way of controlling and teaching the child?
1. Does the mother threaten punishment to control the child?
 2. Does the mother seem to believe that parents must instill fear to control the children?
 3. Does the mother believe children should be formed into absolute submission to the parents' wishes?
 4. Does the mother believe you must break the stubborn will of the child to train it properly?
 5. Does the mother see the child's negativism as a rebellion which must be crushed?
25. Does the mother's ideal seem to be a quiet, passive child rather than an active aggressive one?
1. Does the mother strongly disapprove of the child fighting with other children?
 2. Does the mother disapprove of rough active play of the child?
 3. Does the mother try to put a stop to any display of anger or temper by the child?
 4. Does the mother disapprove of any expression of anger directed against herself or the father and suppress it immediately?
26. Does the mother think the child should be free to act independently and be allowed to work or play apart from the parents?
1. Is the mother willing to have the child work alone with no interference by the mother in the test situation?
 2. Does the mother encourage the child to play by itself or without supervision by the mother?

3. Does the mother permit the child to make his own decisions about his activities, where he will go, etc.?
 4. Does the mother seem to be tolerant of separation of the child from herself?
 5. Does the mother seem willing to give the child freedom of action?
27. Does the mother wish to keep the child closely attached to herself?
1. Does the mother continually need to bring the attention of the child back to herself during the testing session?
 2. Does the mother keep the child with herself around the home at all times and in all situations?
 3. Does the mother tend to avoid having anyone else care for the child and seem to be anxious to resume care of the child after the testing session as though the test relationship were an undesirable interpolation?
 4. Are there signs she doesn't feel free to leave the child in the care of others and to be separated from him?
28. Does the mother believe in rigid rules and strict enforcement of those rules?
1. Does the mother seem concerned about not "spoiling" the child, so that she seldom condones or tolerates his behavior?
 2. Does the mother believe that even small children should be able to learn rules of behavior which she sets up?
 3. Does the mother believe that all rules should be strictly enforced?
 4. Does the mother believe in setting rigid limits to the child's freedom of movement about the house or testing room?
 5. Does the mother rigidly enforce restrictions on habits and other behavior?
29. Is she an over-conscientious mother?
1. Does she show concern about a much wider variety of child-rearing problems than most mothers?
 2. Does her concern for her child reflect an exaggerated effort in the mother to prove to herself that she is doing her job?
 3. Did the examiners have the feeling any suggestions they made would be carried out to the letter or even too zealously by the mother?
30. Does this mother have high intelligence?
1. Does she have a wide range of interests?
 2. Is she mentally alert and responsive?
 3. Does her conversation reflect good judgment and understanding?
 4. Does she have an intelligent interest in the study's methods and goals?
31. Does the mother openly express her love and affection for the child?
1. Does the mother often hug or kiss the child?
 2. Does the mother hold the child in her arms to comfort and console him?
 3. Does the mother often smile or speak in soothing tones to the child?
 4. Does the mother praise the child for his behavior or in other ways express her love and approval?
 5. Does the mother seem to express a warm positive affect toward the child during the testing session?
 6. Does the mother immediately respond to any need of the child for attention, care, or sympathy?
32. Does the mother concern herself about the child's achievement?
1. Does the mother seem to be pushing the child to perform tricks or say verses?
 2. Does the mother seem to demand more achievements than the child easily attains?
 3. Does she teach the child to do things he can't do in the testing situation?
 4. Does she urge the child to perform in mental tests?
-

APPENDIX B

MATERNAL BEHAVIOR RESEARCH INSTRUMENT:
Form II. Quantification of Interview Data

-
1. Does this mother ignore or reject her child?
 1. Does she comment on how much work or trouble the child is?
 2. Would she be willing to have others assume most of the responsibility for the care of the child?
 3. Does the mother seem to know very little about the child?
 4. Does she tend to overlook the needs of the child?
 5. Does she give the impression that the child is not her principal interest?
 2. Does this mother seem cooperative overtly? (Exclude officiousness, interventions, self-abasing cooperation.)
 1. Does this mother seem to share the interviewing situation with the interviewer?
 2. Is she considerate of the needs of others?
 3. Is her behavior with her husband, relatives, and friends indicative of empathy with their problems?
 4. Does she accept the goals of the interview understandingly?
 5. Does she seem to be able to give and take in interaction with others?
 6. Does she have informal, friendly working relationships with members of the family?
 7. Does she enjoy cooperating with others?
 3. Is she a dependent mother?
 1. Does she demand a great deal of help from servants, her family, and friends?
 2. Does she seem unable to handle even minor crises alone?
 3. Does she refrain from making even minor decisions without first asking her husband?
 4. Does this person seek reassurance, sympathy and approval greatly?
 5. Is she unable to tolerate being alone because of her dependency?
 6. Does she seek the care, protection, and attention of others as a reaction to felt weakness?
 7. Is this mother unable to tolerate the disapproval of others?
 8. Does she have an unreasonable need for affection from her children, husband, or others?
 4. Is this mother a sociable person?
 1. Does this mother try to establish a friendly social relationship with the interviewer?
 2. Does she socialize often with neighbors and friends?
 3. Does she have social skills which indicate effort directed toward the establishment of friendly personal relations?
 4. Does she elicit positive personal and social reactions from the interviewer?
 5. Does this mother tend to have many close friends with whom she participates in group activities, visiting, clubs, etc.?
 6. Does this mother participate in PTA, clubs, church groups, or other social group activities?
 5. Does this mother have a positive evaluation of the child?
 1. Does the mother think the child is attractive in appearance?
 2. Does the mother think the child is intelligent in ways she approves?
 3. Does the mother think the child has a desirable emotional disposition?
 4. Does the mother typically approve of the child's behavior?
 5. Does the mother think of the child as sociable with adults and other children in ways she herself values?
 6. Is there an absence of fault finding in the mother's behavior with the child?

6. Does this mother seem to reject the role of homemaker?
 1. Does she complain about the tension or boredom of household chores?
 2. Does she seem to lack enjoyment of her duties as a housewife and mother?
 3. Does she use work or social activities as a way of escaping from her home-making role?
 4. Does she give the impression that she feels held down or shut-up in the home?
 5. Does she seem to value other roles more or be more interested in activities other than that of mother and housewife?
7. Does this mother tend to be irritable?
 1. Does she talk about her frustration and irritation with the child's behavior?
 2. Is she abrupt and rough in her handling of the child or is her voice harsh and unsympathetic?
 3. Is her typical interaction with the child harsh or critical?
 4. Does the interviewer feel that the mother is irritable or hostile?
 5. Is this mother sarcastic, critical, antagonistic, prone to antipathies, quarrelsome, etc.?
8. Does this mother tend to keep her child socially isolated?
 1. Is the child kept away from adults other than the parents?
 2. Is she reluctant to have the child play with other children?
 3. Does she tend to discourage social activities outside the family circle?
 4. Does she consciously or by neglect limit the child's social contact with school mates, neighborhood friends, etc.?
 5. Does the mother discourage dating, parties, or other heterosexual social contacts?
9. Does the mother seem to be unaware of the fact that a child has a mind of his own and that he should be able to do his own thinking without forceful intrusion by the mother?
 1. Does the mother seem to be training the child to think of himself as the mother's possession?
 2. Does the mother seem to think she should be able to direct what he will think or say at any time?
 3. Does the mother's approval of the child seem to be strictly dependent upon the degree to which he adapts his thinking and actions to her wishes?
 4. Does the mother seem to think that she should know what the child is thinking and doing at all times?
10. Does the mother relate to the child as an equal?
 1. Does the mother participate in the child's activities in an informal and unintrusive manner?
 2. Does the mother talk with the child about his interests, concerns, and enthusiasms in a helpful way?
 3. Does the mother tend to reduce emphasis upon age and role differences between mother and child?
 4. Does the mother enjoy spending time with the child?
11. Does this mother have a good marital relationship?
 1. Does she have a close, affectionate relationship with her husband?
 2. Does the mother have a great deal of pride and respect for her husband?
 3. Do the mother and father enjoy activities together?
 4. Does the mother uncritically accept her husband?
 5. Is the mother willing to talk about her husband in a positive way?
 6. Do the mother and father agree on family decisions?
 7. Do the mother and father share an interest in the home, work, ideas, play, etc.?
 8. Does the mother seem to enjoy her relationship with her husband?

12. Does the mother perceive the child as more a burden and inconvenience than a source of happiness?
 1. Does the mother talk about the tension and anxiety of child-rearing?
 2. Does the mother see the child as interfering with other activities and pleasures?
 3. Does the mother feel caring for the child is a burden rather than a joy?
 4. Does the mother talk primarily about problems she has in caring for the child?
 5. Does the mother talk about the child's deficiencies rather than his assets or achievements?
13. Does this mother tend to baby her child or foster dependency in him?
 1. Does she often do things for the child he could do for himself?
 2. Does the mother attempt to protect the child from difficult or demanding situations?
 3. Does the mother see the child as weak, helpless, immature and as needing excessive help, care, and support?
 4. Does the mother try to protect or shelter the child from school work or school authorities?
 5. Does the mother tend to anticipate the child's needs excessively and shelter it from even normal effort?
14. Does the mother appear overtly anxious?
 1. Does she seem tense and inhibited during the interview?
 2. Is there a disruption or prevention of activity of the mother due to her anxiety?
 3. Is the mother easily upset by accidents, difficulties, illness or day-to-day events?
 4. Is she fearful of social situations, injury, the dark, or generally apprehensive?
 5. Does the mother seem to show vigilance, restlessness and inability to relax?
15. Does the mother seem punitive and unkind?
 1. Does the mother seem to lack sympathy for the problems of the child?
 2. Does she seem sarcastic toward the interview?
 3. Does she talk in a negative vein about friends, husband, school, or associates?
 4. Does she tend to elicit hostility on the part of the interviewer?
 5. Does she describe the child critically and unsympathetically?
 6. Does the mother treat the child like a stranger, or with hostility?
16. Does the mother tend toward positive emotional states?
 1. Is she cheerful?
 2. Is she happy?
 3. Is she contented?
 4. Is she enthusiastic?
 5. Does she take pleasure in her activities?
 6. Is she often smiling or laughing?
 7. Does she have a sense of humor?
17. Extent and intensity of emotional and behavioral involvement with the child?
 1. Is the mother very much interested and involved in the child's behavior and achievements?
 2. Does the mother have a tendency toward intense emotional involvement in her interactions with the child?
 3. Does the mother spend a great amount of her time with the child?
 4. Is the mother's primary cathexis with the child?
 5. Does the mother have intense expectations concerning the child?
 6. Does the mother have intense needs in her relationship with the child?

18. Does the mother attempt to use fear as a way of controlling and teaching the child?
 1. Does the mother threaten punishment to control the child?
 2. Does the mother believe children should be formed into absolute submission to the parents' wishes?
 3. Does the mother seem to believe that parents must instill fear to control children?
 4. Does the mother attempt to break the will of the child and impose her own will upon him?
 5. Does the mother try to be an absolute, restrictive authority by use of fear or severe penalties?
19. Does the mother think the child should be free to act independently and should be allowed to work or play apart from his parents?
 1. Is the mother willing to have the child work alone without interference from the parents?
 2. Does the mother encourage the child to pursue his own interests without supervision by the mother?
 3. Does the mother permit the child to make his own decisions about his activities?
 4. Does the mother seem to be tolerant of the child's independent behavior apart from the home and herself?
 5. Does the mother seem willing to give the child freedom of action?
20. Does the mother wish to keep the child closely attached to herself?
 1. Does the mother try to keep the child's activities within the home?
 2. Does she tend to avoid letting the child have a close relationship with others?
 3. Does she seem unable to tolerate separation from the child?
 4. Does the mother seem to be preventing independent behavior which would take the child outside the home?
 5. Does the mother attempt to keep the social life of the child limited to family activities?
21. Does the mother believe in rigid rules and strict enforcement of those rules?
 1. Does the mother seem concerned about "spoiling" the child, so that she seldom condones or tolerates his behavior?
 2. Does the mother believe in setting up numerous rules and regulations?
 3. Does the mother believe in strict enforcement of rules?
 4. Does the mother believe in setting rigid limits on the child's freedom of expression and movement?
22. Does this mother have high intelligence?
 1. Does she have a wide range of interests?
 2. Does she have a fluent use of language?
 3. Is she mentally alert and responsive?
 4. Does her conversation reflect good judgment and understanding?
 5. Is she interested in current events, literature, politics, art, or other intellectual topics?
 6. Does she use reading, classes, or other informed sources to learn more about child-rearing?
23. Does the mother openly express her love and affection for the child?
 1. Does the mother express her affection for the child?
 2. Does she praise the child and is she proud of his behavior?
 3. Does the mother respond to the needs of the child?
 4. Does the mother seem to be interested in the happiness and well-being of the child?
 5. Does the mother express her enjoyment of the child?

24. Does the mother set high standards of achievement for the child?
 1. Does the mother seem interested or concerned about the child's school achievement?
 2. Does the mother seem intent upon the child's social success?
 3. Does the mother set high standards for the child in his order, neatness, and cleanliness?
 4. Does the mother seem to demand more achievement than the child easily attains?
 5. Does the mother set high goals, try to instill in the child a high level of aspiration, and spur the child on to further achievement?
 6. Does the mother have high educational or occupational goals for the child?
 25. Does this mother communicate freely in the interview?
 1. Does she talk freely about her relationship with the child?
 2. Does she talk freely about her family, social activities, interests?
 3. Does the observer feel that the mother is uninhibited and unguarded in her conversation?
 4. Does the mother offer information that would be relevant or does she give it reluctantly?
 5. Does the mother cooperate by making her time available and by freely responding to inquiry?
 26. Does this mother withdraw from external involvements?
 1. Does she seem to be preoccupied with internal processes?
 2. Does she seem to lack a capacity for genuine warmth and affection?
 3. Does she tend to be isolated and seclusive?
 4. Does she avoid involvement with others and have a need for privacy?
 5. Does she live in a world of books, daydreams, music or other asocial pursuits?
 6. Are her contacts with others impersonal and distant?
 27. Positiveness of mother child relationship? (Differentiate from over-possessiveness.)
 1. Does the mother experience and express affection for the child?
 2. Is the mother capable of empathy with the needs and concern of the child?
 3. Does the mother take an interest in the child's social life, hobbies, sports, etc.?
 4. Does the mother enjoy an informal friendly relationship with the child?
 5. Is the child a source of enjoyment for the mother?
 6. Do the mother's activities indicate an orientation toward the child's welfare rather than toward her own needs?
 7. Does the mother trust, respect, and communicate freely with the child?
 28. Does the mother seem to wish to control the child?
 1. Does the mother wish to control the child's activities?
 2. Does she wish to have absolute, unquestioned authority over the child?
 3. Does the mother thwart the attempts of the child at self-assertion?
 4. Does she establish strict and restrictive regulations?
 5. Does she give the child a minimum of freedom to determine his own activities, friends, etc.?
-

A COMPARISON OF NORMAL AND IRRADIATED MONKEYS ON AN ODDITY-REVERSAL PROBLEM*

Department of Psychology, University of Texas

A. A. McDOWELL AND W. LYNN BROWN¹

A. PROBLEM

Settlage, Zable, and Harlow (10) have found that the utilization by monkeys of the same stimulus cues in antagonistic response patterns is a sensitive indicator of brain lesions. The performance of prefrontal animals on tests of this nature was consistently poorer than that of normal controls. Urner (11) reports that rats exposed to 400 *r* gamma radiation suffer a decrement in the reorganization of a response series, learned before exposure, into a new response series, after exposure. This decrement was, however, transient.

The present study was undertaken to determine if monkeys show a similar decrement as a latent effect of whole-body irradiation. The manifestation of such a decrement in the chronic irradiated monkey would be suggestive of degeneration of neural tissue as a consequence of radiation exposure.

B. METHODS

1. Subjects

Thirty-five male rhesus (*Macaca mulatta*) monkeys were employed as subjects. These subjects were grouped into three subgroups: a control group of eight animals, a low-dose group consisting of 13 animals, and a high-dose group consisting of 13 animals. The dose range for the animals in the low-dose group was between 10 to 16 neutron rep plus 70 to 140 gamma roentgens, and the dose range received by the animals in the high-dose group was between 27 to 54 neutron rep plus 284 to 557 gamma roentgens, the source including both gamma and neutron irradiation. The exposure of the animals of the experimental groups to radiation preceded the present study by two years.

* Received in the Editorial Office on April 25, 1958.

¹ This work was conducted at the Radiobiological Laboratory of The University of Texas and the United States Air Force, Austin, Texas, and was supported (in part) with funds provided under Contract AF 41(657)-149 and the School of Aviation Medicine, USAF.

2. Apparatus

All training of animals in this experiment was conducted in a modified version of the Wisconsin General Test Apparatus, the animal's carrying cage serving as the holding cage for the test procedure. A stimulus tray holding three foodwells, spaced six inches apart center-to-center, was used. The problem utilized the same three objects on both original and reversal training. These objects were a circular red painted wood block, a circular green painted wood block, and a triangular red painted wood block.

3. Procedure

In original training, response to the object which was odd in color was rewarded and each animal was trained to the criterion of two successive days with two or less errors per day. Twenty-four test trials were given each day. In reversal training, response to the object which was odd in form was rewarded and each animal was tested to the same criterion as in the original training.

The number of errors to criterion made by each monkey on both the original training and the reversal training was recorded and the results were arranged for analysis. Comparisons between errors to criterion on original training and errors to criterion on reversal training were made for each group using Wilcoxon's paired replicates test (9). The H Test (3) was used to compare the groups with respect to original training errors to criterion, reversal training errors to criterion, and negative savings scores. Each *negative savings score* was derived by dividing the subject's errors to criterion on reversal training by his errors to criterion on original training.

C. RESULTS

Figure 1 shows the mean errors per group for successive sixths of learning to criterion on both the original and reversal training. Statistical comparison of the three groups on errors to criterion in original learning yielded no significant difference ($H = 2.45$, $p = 0.30$). A similar comparison for reversal learning also yielded no significant difference ($H = 1.47$, $p = 0.50$). The increase in errors to criterion during reversal learning over errors to criterion during original learning was significant beyond the 0.01 level of confidence for each group.

The mean negative savings scores for each group are shown in Figure 2. Statistical comparison of the three groups with respect to negative savings scores gave a significant difference ($H = 6.62$, $p = 0.05$). The control

group animals showed the least savings and the high-dose irradiated animals showed the greatest savings.

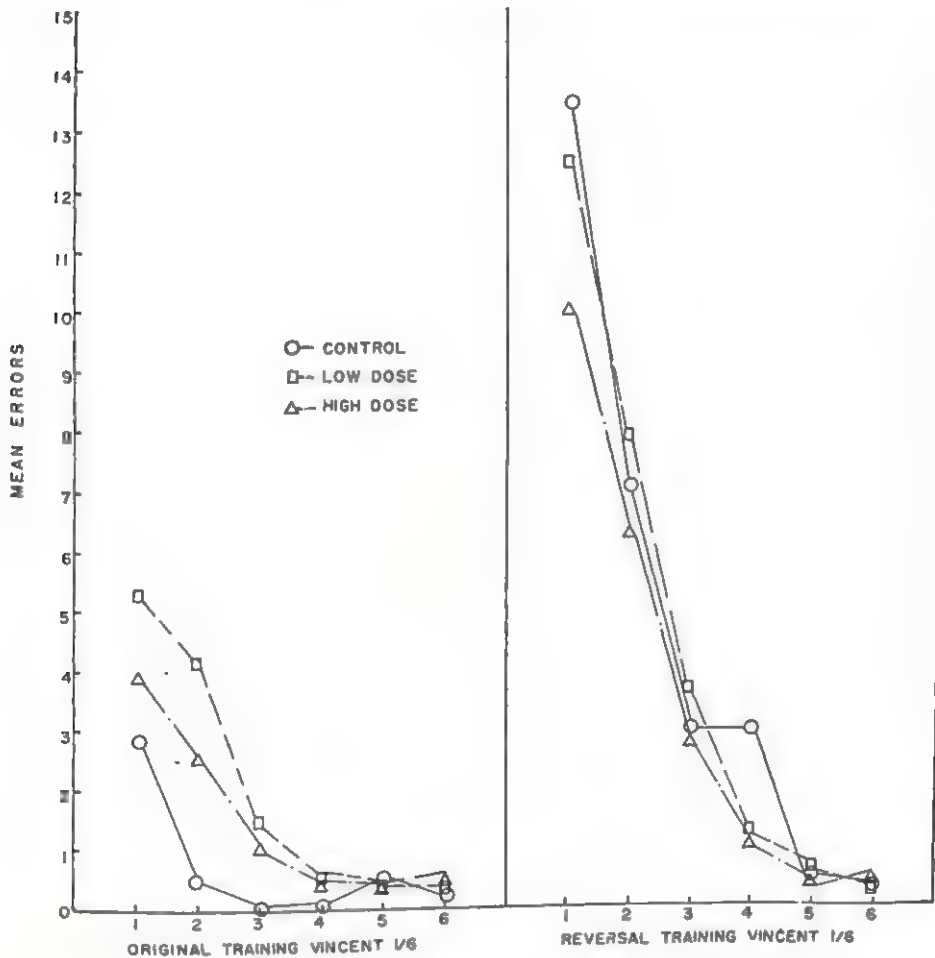


FIGURE 1

MEAN ERRORS FOR SUCCESSIVE SIXTHS OF LEARNING TO CRITERION ON ORIGINAL TRAINING AND ON REVERSAL TRAINING FOR CONTROL, LOW-DOSE IRRADIATED, AND HIGH-DOSE IRRADIATED MONKEYS

D. DISCUSSION

The results of the present study demonstrate that the chronic irradiated male monkey (within the dose range used) utilizes the same stimulus cues in antagonistic response patterns quite as efficiently as the normal monkey.

The irradiated animals, in fact, show significantly smaller negative savings scores than their normal controls. This suggests a facilitation accruing to the irradiated animal with respect to problems of this type.

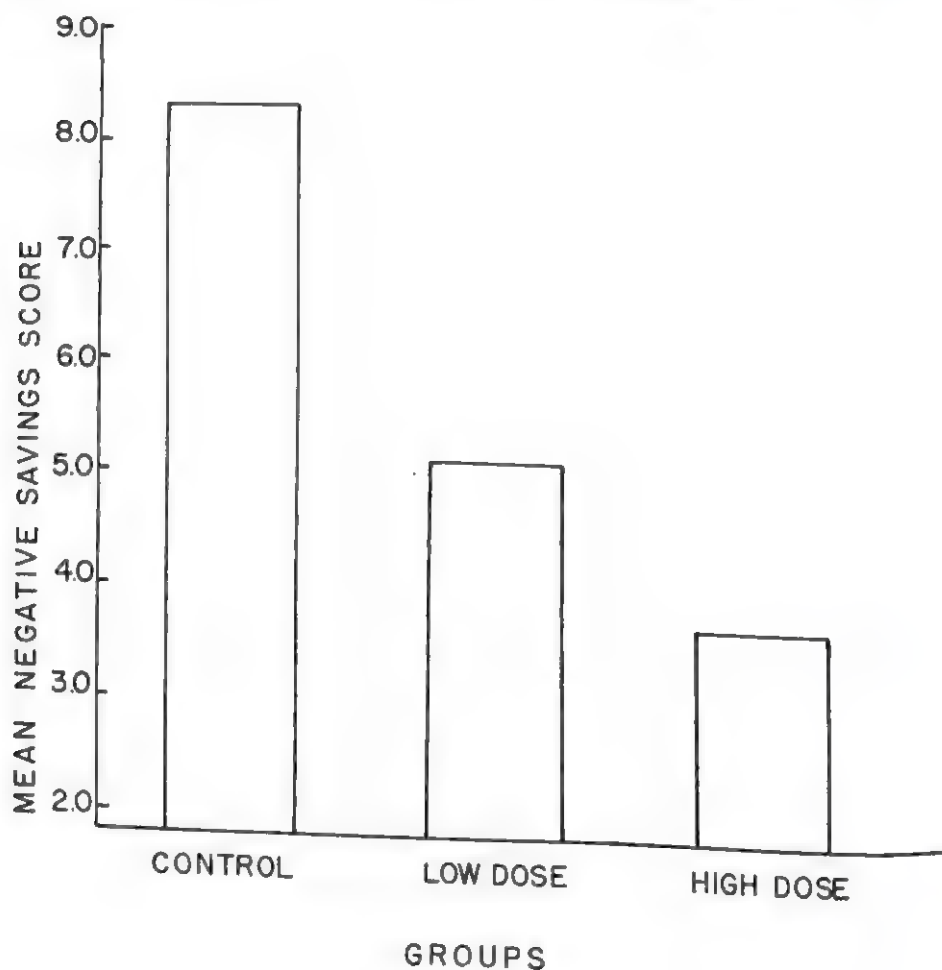


FIGURE 2

MEAN NEGATIVE SAVINGS SCORES FOR CONTROL, LOW-DOSE IRRADIATED, AND HIGH-DOSE IRRADIATED MONKEYS

These findings are in sharp contrast to those of Warren, *et al.* (12) who reported a significantly superior performance by normal monkeys over irradiated monkeys on the solution of discrimination reversal problems. This contrast in findings may be due to differences in experimental design. It

may, however, reflect an artifact of data treatment. Recent research by McDowell and Brown (7) indicates that the irradiated monkey, after having learned a response, suffers less interference from irrelevant stimuli than does the normal monkey on future responses of the same sort. This suggests that the probability of errors on the first trial of all reversal problems in the study by Warren and co-workers would be greater for the irradiated animals than for the normal animals. If this were the case, the inclusion of first trial reversal data in the group comparisons could conceivably account for the reported superiority of performance of the normal animals.

The finding of performance facilitation in the irradiated animal is in accord with the suggestions of Harlow and Moon (4), and of Riopelle, Grodsky, and Ades (8), working with monkeys, and of Blair and Arnold (1) working with rats, that irradiation may facilitate learning task performance. McDowell (5) has demonstrated chronic decreased distractibility in the irradiated monkey, and McDowell and Brown (6, 7), on this basis, have specifically hypothesized and demonstrated a performance facilitation in the chronic irradiated monkey on discrimination problems with reduced stimulus cues and on the spatial delayed response problem. Brown, Carr, and Overall (2) have reported the failure of irradiated monkeys to associate peripherally placed stimuli in contrast to control monkeys who form such associations. All of the above results support a hypothesis of an increase in response thresholds to extraneous stimuli with a consequent narrowing of attention in the irradiated monkey.

E. SUMMARY

Eight normal, 13 low-dose irradiated, and 13 high-dose irradiated monkeys were tested on an oddity-reversal problem which required the utilization of the same stimulus cues in antagonistic response patterns for correct solution. In the original training, each animal was tested 24 trials a day to the criterion of two successive days with two or less errors per day on response to the object which was odd in color. In reversal training, each animal was tested to the same criterion on response to the object which was odd in form.

The following results were obtained:

1. No consistent differences were observed in the number of errors recorded by the three groups to reach either the pre- or postreversal criterion.
2. All groups showed a statistically significant increase in errors to criterion on reversal learning over errors to criterion on original learning.
3. The groups showed a statistically significant difference in negative

saving scores indicating a superiority of the irradiated animal over the normal animal with respect to reversal problems of this type.

REFERENCES

1. BLAIR, W. C., & ARNOLD, W. J. The effects of cranial X-radiation and retention of maze learning in rats. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 525-528.
2. BROWN, W. L., CARR, R. M., & OVERALL, J. E. The effects of whole-body radiation upon association of peripheral cues. School of Aviation Medicine, USAF Report No. 47, 1958.
3. EDWARDS, A. L. *Statistical Methods for the Behavioral Sciences*. New York: Rinehart, 1954.
4. HARLOW, H. F., & MOON, L. E. The effects of repeated doses of total-body X-radiation on motivation and learning in rhesus monkeys. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 60-65.
5. McDOWELL, A. A. Comparisons of distractibility in irradiated and nonirradiated monkeys. *J. Genet. Psychol.*, 1958, **93**, 63-72.
6. McDOWELL, A. A., & BROWN, W. L. Facilitative effects of irradiation on performance of monkeys on discrimination problems with reduced stimulus cues. *J. Genet. Psychol.*, 1958, **93**, 73-78.
7. ———. Latent effects of chronic whole-body irradiation upon the performance of monkeys on the spatial delayed-response problem. *J. Gen. Psychol.*, 1959, **61**, 61-64.
8. RIOPELLE, A. J., GRODSKY, M. A., & ADES, H. W. Learned performance of monkeys after single and repeated X-radiation. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 521-524.
9. SEIGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
10. SETTLAGE, P. H., ZABLE, M., & HARLOW, H. F. Problem solution by monkeys following bilateral removal of the prefrontal areas: VI. Performance on tests requiring contradictory reactions to similar and to identical stimuli. *J. Exper. Psychol.*, 1948, **38**, 50-65.
11. URMER, A. H. The effect of gamma radiation on the reorganization of a learned complex maze habit. Unpublished doctor's dissertation, Univ. Texas, 1957.
12. WARREN, J. M., KAPLAN, S. J., & GREENWOOD, D. D. The solution of discrimination reversal problems by normal and irradiated monkeys. School of Aviation Medicine, USAF, Report No. 16, 1955.

Department of Psychology
University of Texas
Austin 12, Texas

RELATIONSHIP OF SOCIOECONOMIC STATUS TO CHILDREN'S OCCUPATIONAL ATTITUDES AND INTERESTS*

School of Education, University of California, Berkeley

LAWRENCE H. STEWART¹

A. INTRODUCTION

One of the striking ways in which adults differ from one another is in their attitudes toward occupations. In so far as these attitudes are reflected in scores on vocational interest inventories, they remain remarkably stable after late adolescence. That these attitudes are significant in the process of occupational choice and in occupational adjustment has been conclusively demonstrated (2, 22, 23, 24). Yet, there are many aspects of occupational attitudes about which very little is known. When, for example, do these attitudes develop? What factors are related to the development of different types of attitudes? Questions such as these are largely unanswerable on the basis of existing research.

A number of studies (7, 32) have pointed to the waste of manpower potential which occurs because many high-ability students do not obtain the level of education necessary to enter occupations in which they could function at a level consonant with their ability. If Ginzburg (6) was correct in his interpretation of the findings of the Kinsey report, discrepancies between level of aspiration and abilities are closely related to social class values. Ginzburg suggested on the basis of Kinsey's findings that a socially mobile individual adopts the value system of the social class toward which he is moving, long before changes in his behavior can be observed, and high level of aspiration for education and vocational achievement appears to be characteristic of middle class values. Thus, if one major concern of education and of counseling is to implement maximum utilization of each individual's talents for the welfare of society, knowledge of attitudes—how they develop and how they can be changed—is fundamental. Such knowledge is equally important to the counselor who is concerned primarily with the happiness and general welfare of the individual.

One source of this knowledge should be research on children's interests.

* Received in the Editorial Office on May 5, 1958.

¹ This research was supported by a grant from the Research Committee, College of Education, University of California, Berkeley.

Examination of studies in this area indicates, however, that typically they have been normative surveys (e.g., 12, 13, 16). In general, findings of these surveys suggest a recapitulation theory of interest development—that children pass in a gradual and predictable manner through various stages, each of which has characteristic interests.

A major difficulty with normative studies is that they often provide no means by which variation from the norm can be explained or even studied. The surveys indicate differences in the interests along broad dimensions such as sex, race, and socioeconomic status, but even here similarities are more apparent than differences. A more meaningful type of investigation might attend more to deviations from the dominant sub-groups. What is the significance, for example, of the fact that a boy's interests are more like those of girls than like those of boys of his own age group? What are the characteristics of a boy from a high socioeconomic status background who has interests similar to boys from lower status homes? In the long run, studies of deviations such as these may provide basic keys to the dynamic interaction of factors in attitude development.

A second source of relevant information should be theories of vocational interest development. A number of such theories have been proposed. Reviews of theories of vocational interest are amply available elsewhere (2, 24). These theoretical formulations have been, for the most part, enumerations of factors presumably operative in early childhood. Support for the theories are primarily extrapolations of findings from studies of older adolescents and adults. As far as the writer has been able to determine, Tyler's longitudinal study of children's interests (26, 27) is the only one designed to investigate directly factors associated with the origins of vocational interests. It is still too soon to determine whether her measures of children's interests are related to measures of adult vocational interests such as the Strong Vocational Interest Blank. Neither normative surveys nor theories of interest provide a sufficient description of how various patterns of occupational attitudes develop. As more research data accumulate, more adequate theories can be formulated. Findings of studies such as the investigation reported here contribute to that end.

1. Socioeconomic Status and Attitudes

Bordin (1) has hypothesized that in responding to the items on the Strong Interest Blank an individual is expressing a particular concept of himself in terms of occupational stereotypes. Tyler (26) suggested that the individual expresses the concept of himself in terms of the stereotypes which the

method of measurement makes available to him. She presented evidence which indicates that the stereotypes which a measure of children's interests make available to the individual depend upon factors such as age and sex. One would also suspect that another important factor would be socioeconomic status.

According to Davis (3) socioeconomic status becomes operative almost at birth in terms of how the parents react to the child. The process and goals of acculturation differ markedly for families from different socioeconomic levels. The middle class child is punished (by shame) for physical aggression; he is rewarded (by approval) for being competitive in a more socially approved manner, such as achievement through education. Lower status families, on the other hand, often reward the behavior the middle condemns. Davis has suggested that a person's social reality—and, therefore, his social drives, goals, and values—is determined by the social class to which he belongs.

Studies (e.g., 8, 9, 14, 15) investigating the effects of socioeconomic status on the behavior of adolescents have tended to substantiate, to some degree, Davis' contentions. Socioeconomic status has also been shown to be related to children's interests and activities (15, 16, 18, 21). Apparently the social class in which a child is reared determines, in large measure, his reactions to his environment and, in turn, how others react to him.

2. Problem

This study investigated relationships between certain aspects of socioeconomic status and children's attitudes toward occupations, and relationships between these occupational attitudes and certain types of children's behavior. Specifically, the study was concerned with the following questions:

1. Are perceptions of occupational symbols of social class related to the socioeconomic backgrounds of fifth grade boys?
2. Are their ideas of certain aspects of social class behavior related to the socioeconomic backgrounds of these boys?
3. Are the reputations of fifth grade boys among their peers related to their class status?
4. Is there a relationship between children's occupational level interests and their perceptions of social class symbols? Between interests and ideas of social class behavior? Between interests and reputation among peers?

3. Definition of Terms

"Social class" and "socioeconomic status" were used interchangeably in this study. The criterion of social class was father's occupation; the measure

of social class was a revision of Roe's occupational classification scale. "Class levels" refer to the six levels on the Roe scale.

As is evident from the discussion of measures of status by Warner, et al. (30), many criteria of socioeconomic status were not included in this study. Occupation, however, has been demonstrated to be one of the most effective indices of status, and attitude toward occupations was the primary focus of this study. It is with this in mind that such an index of status was used here.

Socioeconomic status is represented by "symbols" which usually call to mind a particular social class when they are mentioned. This study was not concerned whether these symbols accurately portrayed a social class, but whether the subjects perceived them as being related to particular occupations.

B. METHOD

1. General Procedure

In the spring of 1957, the *Dreese-Mooney Interest Inventory for Elementary Grades* (4) was verbally administered to a group of fifth grade boys. Following the inventory, a 50-minute interview was held with each child, during which the subject was asked to associate various symbols of socioeconomic status with a set of nine pictures showing men working at various occupations. A Guess-Who test was then given twice to each boy. The first time he was asked to respond with reference to the occupational pictures; the second time, with reference to his peers.

2. Sample

The sample was composed of 243 fifth grade boys from five elementary schools in one school district located within commuting distance of the San Francisco Bay Area and from one school affiliated with the University of California (Berkeley). These schools represent a variety of types of communities among which are included agricultural, manufacturing, and residential areas.

Finding schools which include children from all social classes was rather difficult because elementary schools tend to be homogeneous as to socioeconomic status since they typically serve only a particular housing development or one section of a community. Each school finally selected included students from a fairly large range of socioeconomic backgrounds. Table 1 indicates the number of subjects from the various status homes.

An effort was made to include all fifth grade boys in the school selected.

This was accomplished in four schools. Illness prevented one or two students in each of the remaining schools from participating.

TABLE 1

Social class (Revised Roe scale)	Number
I	6
II	29
III	32
IV	110
V	54
VI	12
Total	243

Negroes, Orientals, and recent immigrants were excluded from the sample to minimize possible effects of differences in cultural milieu not related to socioeconomic status. Miles (16) found, for example, that the interests of Negro children were quite different from those of white children. Also, girls were not included because studies (16, 21, 26) have clearly shown their interests to be different from those of boys. The sample was limited to the one grade level in order to eliminate the factor of changes in attitudes with age. Fifth grade students were selected for two reasons: their greater ability in communication than is found in younger groups, and the availability of an "occupational level" scale of children's interests which is appropriate for this age group. There is also some evidence (e.g., 21) that somewhere around the ages at which children enter the fourth or fifth grade, status factors begin to assume more importance in their behavior. This, however, remains to be verified by further study.

3. Instruments

a. Drees-Mooney Interest Inventory for Elementary Grades. This instrument is composed of 250 items divided into 10 areas: reading, movies, radio, games and toys, hobbies, things to own, school subjects, people, occupations, and activities. Except for 13 items, the respondent indicates his answer by checking like, indifferent, dislike, or unknown for each item. For the 13 items, the respondent must choose between two activities.

The Drees-Mooney Inventory was scored by means of the "occupational level" scale which has been described in another report (in press). The scale was developed by comparing the responses of 113 sons of business and professional men to those of 107 sons of semiskilled and unskilled workmen

to items on the inventory. The criterion groups were fifth grade boys from schools within a 300-mile radius of the San Francisco Bay Area.

The reliability of the occupational level scale, obtained by means of the split-half procedure, was .776; corrected by the Spearman-Brown formula, .874. This reliability is somewhat higher than those reported on other measures of children's interests (26); it is about equal to those reported on similar measures of adult interests (22). The scale has been shown to differentiate appreciably among boys from different socioeconomic strata as determined by father's occupation. Thus, the reliability and validity of the scale were considered quite adequate for purposes of this study.

b. Stimulus Pictures. Fifty faceless line drawings depicting persons in a variety of occupations were shown to a group of fifth grade students similar to those included in the sample used in this study. On the basis of their responses nine pictures were selected. Two criteria were used. First, the occupations depicted had to be recognizable by all the subjects. Second, it was necessary that there be variation in the social status of the occupations. The pictures finally chosen represented roughly three socioeconomic status groups: upper status represented by a judge, teacher, and scientist; middle level by a mechanic, police sergeant, and salesman; and lower by a laborer, bellboy, and newsboy. A modification of Roe's occupational classification (17) was used to classify these occupations according to level.

Horowitz (10) and Radke et al. (19) have demonstrated in their studies of prejudice that stimulus pictures are effective interview aids in working with elementary school age children; that in discussing the pictures the child reveals many of his feelings which he would not be able to verbalize in abstract terms. Stendler (21) has shown that through reaction to carefully selected pictures, children reveal their awareness and evaluation of social class symbols. These studies suggested that such pictures would provide a satisfactory technique for investigating many of the variables of concern in this research.

The drawings were presented to all subjects in the same order, which was determined by a table of random numbers. All nine pictures were presented at once to the subject with the following instruction: "This is a guessing game. Look at these pictures and see if you can guess about whom I am talking." This was followed by a list of five questions. In each case the subject was asked to justify his answers.

1. Who earns the most money? The least?
2. Which of these men live in the biggest and nicest house? The smallest and poorest?

3. Which of these men is most admired and respected in the community?
A nobody?

4. Which of these men has the most education? The least?

5. Which one would you rather be like?

c. *Guess-If/Who or opinion test.* The items for the Guess-Who test were adapted from Tryon's opinion test (25). The test consists of eight pairs of behavior descriptions, each pair representing two extremes of the same trait or characteristic. Traits which were selected had been shown by previous research to be related to socioeconomic status. The number of traits was limited to eight because the child's fatigue became a factor if the interview exceeded 50 minutes in length. The items are described below in Table 2.

TABLE 2
ITEMS IN GUESS-WHO TEST

1. <i>Fights.</i> This boy enjoys a fight.	<i>Avoids Fights.</i> This boy never fights but lets the other person have his way.
2. <i>Leader.</i> This boy always knows how to start games or suggest something interesting to do so others like to join in.	<i>Follower.</i> This boy waits for somebody else to think of something to do and always likes to follow suggestions which others make.
3. <i>Active in Games.</i> This boy plays games like kickball and dodgeball or likes to run and jump. I'd like to have him on my team.	<i>Sedentary.</i> This boy seldom plays games like kickball and dodgeball. He prefers to read or sit and play quiet games. I would not like to have him on my team.
4. <i>Good Student.</i> This boy is someone who is a very good student and always does well in school work.	<i>Poor Student.</i> This boy is someone who is not interested in school and never does well in his school work.
5. <i>Best Friend.</i> This boy is someone who is my best friend.	<i>Least Liked.</i> This boy is someone I don't like at all; I don't want him for a friend.
6. <i>Friendly.</i> This boy is very friendly. He has a lot of friends and is nice to everybody.	<i>Snob.</i> This boy is someone who is "stuckup." He thinks he is better than anyone else. He thinks that the things he does are best.
7. <i>Like to Visit.</i> This boy is someone whose home we like to visit and like to invite to our home.	<i>Dislike to Visit.</i> This boy is someone we would not visit at home or invite to our house.
8. <i>Protects the Weak.</i> This boy is someone who would never pick on anyone. He sticks up for other children when they are being teased or bullied.	<i>Bully.</i> This boy is someone who is a bully; he often hits and teases girls and smaller children.

The above items were administered twice. The first administration was made with the following instructions: "This is a guessing game. I am going

to describe a boy. See if you can guess *the boys in the fifth grade in your school that are like the person I describe.*"

The second time the Guess-Who test was administered the occupational pictures provided the frame of reference. Instructions were revised as follows: "All of these men in these pictures have sons. See if you can guess *which of these men has a son like the boy I am describing.*"

Tryon (25) has suggested three indications of reliability of the Guess-Who items. The first of these is the concurrence of opinion as observed in the patterns of responses about individuals; i.e., certain subjects seem to be regarded generally as being on one end of a continuum or the other, particularly those who deviate most from the mode. Inspection of responses to the Guess-Who items indicate that the ratings tended to fall into such patterns.

A second indication is the degree of agreement between children who are both judge and subject. A group of 68 boys from one school was divided into two equal groups of raters. The ratings assigned to the 68 boys by one half of the raters were then correlated with the ratings of the other half, thus giving a measure of agreement. The median r_{jj} was .721; corrected by the Spearman-Brown formula, .838. Separate correlations were made for each of the opposite ends of a continuum. This was done when it became apparent that the r 's for extremes were not identical in every case. For example, the r_{jj} for leader was .764; for follower, .265. Most of the r 's for opposite ends of a continuum, however, were of similar magnitude. These findings do not necessarily indicate a high degree of agreement among all judges. Since each individual was rated by only a small number of his peers, the correlations reflect the fact that a small number in each of the two groups saw the individual in the same way. The correlations were probably spuriously high because of the large number of students who received a small number of choices or were not mentioned at all.

A third indication of reliability would be the stability of opinions as determined by test-retest. There was no opportunity of obtaining test-retest information in this study, but other studies (11, 25) have shown that sociometric ratings under similar conditions are about as stable as other frequently used measures of personality variables.

d. *Interview.* All data except the Drees-Mooney Inventory were collected from individual interviews. Thus, in effect, a structured interview was built around the instruments. Collecting data in this manner provided an opportunity to question the subject as to the reasons for his answers and to clarify ambiguous responses. Many seemingly conflicting responses were quite logical after the subject was given an opportunity to explain his answers.

Moreover, it was felt that the interview provided motivation to take the various tasks seriously.

The interviews were conducted by four trained interviewers. One was a psychiatric social worker; the other three were completing requirements for the master's degree in Counseling Psychology. All had had experience in working with children of this age group.

e. Classifying occupations. A modification of Roe's occupational scale (17) was used to classify the fathers' occupations. The revised Roe scale was used because it more nearly places occupations on a socioeconomic continuum than do other scales such as the *Dictionary of Occupational Titles* (D.O.T.). Thus, it is more amenable than the D.O.T. to the type of analysis used in this research.

Roe's scale takes account of interest as well as level of responsibility and training in classifying occupations. Only level, however, was used in this study. The levels are defined in Table 3.

TABLE 3
OCCUPATION LEVELS FROM ROE'S SCALE

I. Professional and Managerial 1
A. High level and varied responsibility
B. Policy making
C. Education at doctoral level or equivalent
II. Professional and Managerial 2
A. Medium level responsibility
B. Policy interpretation
C. Education at or above bachelor level, but below the doctoral level or equivalent
III. Semi-professional and small business
A. Low level responsibility
B. Application of policy
C. Education, high school plus technical school or equivalent
IV. Skilled
V. Semiskilled
VI. Unskilled

C. CLASS DIFFERENCES IN PERCEPTIONS OF SOCIAL CLASS SYMBOLS, IDEAS OF SOCIAL CLASS BEHAVIOR, AND REPUTATION

Stendler (21) has provided evidence from which she concluded that socioeconomic status is clearly related to children's attitudes. From a study of 107 elementary school children from Grades 1, 4, 6, and 8, she found that perception of social class symbols becomes more accurate in the higher grades. Variables such as choice of friends, out-of-school activities, and future occupational goals were related to social class, particularly in the upper grades. Caution is necessary in generalizing from Stendler's data, however. When

her subjects were divided by three status groups, by sex, and by four grade levels, extremely small groups necessarily resulted. Significant findings based on such small *N*'s often do not hold up under cross-validation. The findings presented in this section bear directly on those reported by Stendler.

The questions to be considered are: Are perceptions of occupational symbols of social class related to the socioeconomic background of fifth grade boys? What are the ideas of fifth grade boys about social class behavior, and are these ideas related to their own status? Is the reputation, among peers, of fifth grade boys related to their socioeconomic status?

1. *Perceptions of Class Symbols*

Data relative to perceptions of class symbols were obtained from subjects' reactions to the occupational pictures. It will be recalled that each boy was asked questions such as "Which of these men earn the most money? The least?" Examination of the percentages of subjects ascribing a certain class symbol to a particular occupation indicated that, in general, these boys were aware that social class symbols are related to occupations. Symbols of higher socioeconomic class such as a fine home, high income, importance in the community, and high level of education tended to be ascribed to the professions; a lack of these symbols, to lower level occupations. Whereas most of the perceptions of class symbols were similar to those found in various studies of occupational prestige, two exceptions were observed. "Policeman" was consistently ranked higher than the position typically assigned to him by most occupational classification systems. The boys perceived him to be of equal status to the professions depicted in the drawings in all status characteristics except amount of education. Next to the judge, he was regarded as the most respected member of the community. Comments made in the interviews by the subjects indicated this ranking was made because both judge and policeman were perceived as persons who "control others" or who "keep order." The salesman, on the other hand, was ranked consistently on a status level comparable to that of the unskilled occupations, despite the fact that his clothing clearly placed him in the "white collar" class. This might be explained by the fact that the salesman was depicted as a house-to-house canvasser.

There was a considerable amount of variation in the assignment of class symbols to the various occupations. The question arose as to whether this variation was related to the socioeconomic background of the child. The responses were, therefore, analyzed separately by social class of subjects, who were grouped according to the Roe occupation classification of fathers'

occupations. The percentages of each class assigning a particular class symbol to the nine occupations were then ranked. To the question, "Who has the most money?", for example, Classes I and II ranked scientist first, judge second, and so on for the nine occupations. Responses of the other classes were ranked in a similar manner. Because of small *N*'s in Classes I and VI, they were combined with Classes II and V, respectively.

Relationships between the rankings of various classes were determined by means of the coefficient of concordance (*W*) as described by Walker and Lev (29). This analysis gives an indication of the degree of similarity among the four sets of ranks for a particular class symbol. A high coefficient means that the sets of ranks for the various classes were quite similar. Because of the small number of classes, a correction for continuity was made in each analysis. Socioeconomic group or class was considered as "treatment" or "raters"; the nine occupations, as "subjects."

TABLE 4
SIMILARITY OF PERCEPTIONS OF OCCUPATIONAL CLASS SYMBOLS AMONG BOYS FROM
VARIOUS SOCIAL CLASSES

Question	Coefficient of concordance (<i>W</i>)
Most money	.958
Least money	.900
Biggest house	.964
Poorest house	.939
Most respected	.823
Least respected	.887
Most education	.931
Least education	.910
Who do you want to be like?	.837

The coefficients are shown in Table 4. Their magnitudes indicate agreement far beyond that which might be produced by sampling variance. These findings indicate that the subjects' perceptions of class symbols were relatively independent of socioeconomic status.

Combining rankings for all occupational levels might mask any real differences which exist at particular levels. As a check on this possibility, the data were analyzed again. To the question, "Who has the biggest house?", for example, 94 per cent of subjects from Class III, 84 per cent of Class IV, 80 per cent of Classes I and II, and 76 per cent of Classes V and VI named the judge. These percentages were ranked, Class III receiving a rank of 1; Class IV, 2, etc. Percentages for teacher and scientist were ranked in the same way. The sets of ranks were then analyzed by the analysis-of-variance-by-ranks test as described in Walker and Lev (29). Socioeconomic classes of

subjects were considered as "treatments"; three pictures at a particular level, as "subjects." This procedure was replicated for the three occupations at each of the other two levels. Separate analyses were made for each class symbol.

The analysis-of-variance-by-ranks test, as used here, provides an indication of the extent to which some classes tend to associate a status characteristic with an occupational level more frequently than do other classes. A significant F would mean that the various socioeconomic classes differ in this respect.

In only one instance out of 27 analyses was the obtained F sufficiently large to be considered significant. Boys from lower socioeconomic homes tended to rank the unskilled occupations higher in terms of "living in the nicest house" than did boys from other socioeconomic groups. This relationship, however, was not clear-cut ($.01 < P < .05$). With this one exception, these findings provide further evidence that awareness of class symbols was independent of the socioeconomic status of these fifth grade boys.

Although the coefficient of concordance in Table 4 indicated that expressed occupational preferences of the various classes were quite similar, further analysis of the data shown in Table 5 suggests some tendencies which should

TABLE 5
OCCUPATIONAL PREFERENCE AND SOCIOECONOMIC STATUS*

Social class of subject	Level of occupation		
	Upper	Middle	Lower
I & II	94%	9%	0%
V & VI	63%	39%	4%

* Percentages total more than 100 because some subjects were unable to decide between two occupations. Both choices were counted.

be pointed out. Despite the fact that boys from all classes perceived the policeman as having high status, he was not chosen equally by all classes as a vocational preference. Sons of professional and managerial fathers tended to restrict their preferences to the three upper class occupations. Sons of semiskilled and unskilled workmen, on the other hand, exercised more variation in choice. The latter group still tended to aspire to upper level jobs and not to choose the lower level occupations.

These fifth grade boys were aware of social class symbols but their perceptions certainly were not biased appreciably by their own status. These findings are not in agreement with other studies which have suggested, at least by implication, a substantial socioeconomic bias in children's attitudes (e.g., 21). Do these discrepancies reflect changes in mass media or in the

school curricula during the past decade? Or will the social class of these boys assume more importance in their perceptions as they grow older? These problems are deserving of further study.

2. Socioeconomic Status and Ideas of Social Class Behavior

Do fifth grade boys relate certain types of children's behavior to socioeconomic status as determined by fathers' occupations, and, if so, are these stereotypes related to the subjects' social backgrounds? Data relative to these two problems were obtained from the Guess-Who test administered in connection with the occupational pictures. Examination of the responses to the Guess-Who test indicates that sons of professional and managerial fathers were perceived as tidy, friendly, avoiding fights, leaders, good students, best friends, attractive to visit, and protectors of the weak. Sons of lower status workmen were more often perceived as evidencing opposite types of behavior. The only trait which did not appear to differentiate clearly among the occupational levels was active-sedentary. Ratings on this trait were distributed evenly among all nine occupations.

Sons of the policeman and of the salesman were perceived differently from the trends discussed above. The policeman's son was perceived as being quite similar to sons of men in higher status occupations. The salesman's son, on the other hand, was perceived as being similar to sons of unskilled workmen.

To obtain some indication of the relationship between socioeconomic background of the subjects and their perception of the behavior stereotypes, the ratings of nine occupations by subjects from each class were ranked for each trait in the test. For the trait "tidy," for example, subjects from Classes I and II ranked sons of teachers first, judges second, and so on. This procedure was repeated for each of the other classes. The similarity of the rankings by the various classes was determined by the coefficient of concordance. An analysis was made for each trait. Social class was considered as "raters" or "treatments"; occupations, as "subjects." Correction for continuity was made because of the small number of raters.

The results of these analyses are shown in Table 6. It is apparent that children from various classes held remarkably similar ideas of children's social class behavior. The sets of ranks for all questions showed agreement far beyond that expected as a result of sampling error.

Because of the possibility that including all levels of occupations in the same analyses might obscure some of the relationships between social class and ideas about social class behavior, analysis-of-variance-by-ranks tests were

applied for each of the three levels of occupations depicted in the pictures. For the trait "tidy," for example, 74 per cent of socioeconomic Classes I and II named the judge's son, 81 per cent of Class III, 80 per cent of Class IV, and 67 per cent of Classes V and VI. These percentages were ranked, Class III receiving a rank of 1, IV, a rank of 2, etc. This procedure was duplicated for each occupation at each level. A separate analysis was made for each trait. Status groups were considered as "treatments"; the three occupations at any level, as "subjects."

TABLE 6
SIMILARITIES OF IDEAS OF SOCIAL CLASS BEHAVIOR AMONG BOYS FROM VARIOUS SOCIAL CLASSES

Stereotype	Coefficient of concordance (W)
Tidy	.912
Untidy	.957
No fight	.925
Likes to fight	.925
Leader	.875
Follower	.871
Active	.768
Sedentary	.616
Good student	.862
Poor student	.878
Best friend	.930
Not best friend	.895
Friendly	.831
Unfriendly	.678
Like to visit	.914
Don't like to visit	.910
Protect the weak	.918
Bully	.914

These analyses indicated the following significant differences among socioeconomic classes:

1. Subjects from Classes I, II, and III were more likely to see sons of the teacher, judge, and scientist as leaders than were subjects from other socioeconomic classes ($.01 < P < .05$).
2. Subjects from Classes I, II, and IV were more likely to perceive boys of middle status parents as being followers or non-leaders than were the subjects from other classes ($P < .01$).
3. A larger percentage of subjects from Class III indicated they would like to visit in the homes of the unskilled workmen than did subjects from Classes I, II, V, and VI ($P < .01$).
4. More subjects from Classes I and II tended to see sons of upper status

homes as being concerned with the welfare of the underdog, i.e., "protects the weak," than did those from Classes III, V, and VI ($.01 < P < .05$).

5. Subjects from Class III were more likely to see sons from lower status homes as being protectors of the weak than were subjects from Classes I, II, and IV ($.01 < P < .05$).

6. Subjects from Classes I and II tended to designate more often the men in the pictures depicting the lower level occupations as having sons whom they would least like to have as a friend ($.01 < P < .05$).

7. Subjects from Classes I, II, and IV tended more often to perceive the men in the lower status occupations as having sons who were unfriendly ($P < .01$).

When subjects from the different social classes disagreed, it was most often relative to their ideas about the behavior of children from upper and lower status homes. Although out of a possible 48 there were only seven significant *F* ratios, four of which were not clear-cut, the number obtained is more than would be expected by chance.

The most striking finding, however, was the amount of agreement among subjects of the different socioeconomic levels. This further substantiates the previous finding that perceptions of class symbols were not greatly biased by the subjects' class background.

3. *Reputation and Socioeconomic Status*

Do subjects from different social classes rate their peers in the same way; e.g., do upper status boys rate other upper status boys in the same way boys from other classes rate them? Are peer ratings related to the socioeconomic backgrounds of the children being rated? Each of these questions will be discussed in turn.

For the peer group analysis, the classes were combined into three levels: upper, including Classes I and II; middle, III and IV; lower, V and VI. Ratings for each level of peers were grouped according to the class of the rater. In order to avoid overlapping of status groups, Class IV raters were omitted from this analysis although they were included in those being rated. The next step was to correlate the ratings of each social class peer by raters from his own class with the ratings of the same peer group by raters from the other social classes.

The results which are shown in Table 7 read as follows: The correlation between Class I ratings of upper level peers and Class II ratings of the same group was .893; Class I and Class III, .906, etc. The magnitude of the intercorrelations indicates that reputation of peers was relatively independent of the socioeconomic status of the raters.

To determine the relationship between reputation and socioeconomic status of the boys being rated, the sociometric ratings were analyzed by means of analysis of variance. For these analyses, the traits, e.g., "tidy-untidy," were combined. A separate analysis was made for each continuum. None of the *F* ratios obtained in this application were significant. Thus, these findings indicate that the reputations of these boys were independent of the socioeconomic status of both the person being rated and the rater. Factors related to reputation among peers remain to be determined.

TABLE 7
INTERCORRELATIONS OF SELECTED SOCIOMETRIC RATINGS OF PEER GROUPS
(Grouped According to Socioeconomic Status of Rater and of Peers)

Social class of rater	Social class of peer	II			III			V			VI		
		U	M	L	U	M	L	U	M	L	U	M	L
I	Upper N = 35	.893			.906			.869			.947		
	Middle N = 142		.978			.997			.977			.996	
	Lower N = 66			.913			.950			.932			.992
					.917			.807			.941		
II	Upper					.985			.952			.983	
	Middle						.955			.925			.897
III	Upper							.934			.948		
	Middle								.985			.991	
V	Upper									.982			.944
	Middle										.863		
	Lower											.974	
													.921

4. Summary

Findings presented in this section have been concerned with relationships between socioeconomic status and perceptions of occupational class symbols, between status and ideas of social class behavior, and between status and reputation of fifth grade boys. These boys had rather realistic perceptions of the relationships between occupations and symbols of social class in so far as they tended to agree with most studies of occupational prestige. They also had formed definite ideas of social class behavior. While there was a small number of significant relationships, most of the data indicated that class symbols were independent of raters' socioeconomic status backgrounds. Furthermore, peer ratings were not related to the class status of the rater or to that of the person being rated.

These findings support those of Stendler in that the children perceived a relationship between symbols of social class and occupation. They did not support her findings, however, in the matter of the relationship of socioeconomic status and perceptions of status symbols or behavior as determined by the sociometric ratings, which were here found to be independent.

D. RELATIONSHIP OF "OCCUPATIONAL LEVEL" INTERESTS TO STATUS VARIABLES AND TO REPUTATION OF PEERS

What is the significance of deviations from typical social class status perceptions and of interest patterns? Findings reported in the previous section were discussed as group trends, but there was a considerable amount of variation among subjects at all class levels. Moreover, while there were significant differences in the mean occupational level scale scores for the social classes, the distributions overlapped to an appreciable degree. (Findings relative to occupational level scale refer to study now in press.)

As shown in the previous section, deviations from the class mean in ability to ascribe class symbols to occupations and in information about certain aspects of peer behavior were independent of social class status. These deviations were also independent of the variations in occupational level interest scores. Subjects with similar interest scores seemed to perceive and evaluate social class symbols in a similar manner,—that is, without regard to their own socioeconomic status. This section will be devoted to findings which support these observations.

In analyzing the data relative to perception of class symbols, subjects were classified according to status and interest scores in the following manner: subjects from Classes I, II, and III were combined into a high status group; Classes V and VI, into a low status group. Class IV subjects were omitted from the analysis in order to avoid overlapping of the status groups. The high and low status groups were further divided into high and low occupational level interest groups. An interest score was considered high if it fell at or above the mean of the entire sample; low, if it fell below the mean. The number of subjects in each of these groups is shown in Table 8. Inspection of the percentages of each status-interest group who ascribed a

TABLE 8

	N
High status-high interest	36
High status-low interest	31
Low status-high interest	24
Low status-low interest	42

particular status characteristic to each occupation indicated that the groups were remarkably similar in their ratings.

As a more direct indication of the relationships between the perceptions of the status-interest groups, rank-order correlations were computed among the ratings of the groups for each class symbol. The correlations are shown in Table 9. There was no pattern among these correlations. From these

TABLE 9
PERCEPTIONS OF OCCUPATIONAL CLASS SYMBOLS AND OCCUPATIONAL LEVEL INTEREST
SCORES RANK-ORDER CORRELATIONS BETWEEN INTEREST-STATUS GROUPS

Interest	Status	Most money	Least money	Nicest house	Poorest house	Most re-spected	Least re-spected	Most edu-cation	Least edu-cation	Want to be like
High	High	.910	.946	.975	1.00	.667	.908	.975	.979	.392
Low	High									
High	Low									
High	High	.921	.971	.862	.929	.892	.921	.958	.996	.675
High	High									
Low	Low	.934	.933	.962	.938	.721	.792	.983	.971	.558
Low	High									
High	Low	.954	.992	.900	.929	.712	.838	.983	.983	.812
Low	High									
Low	Low	.946	.908	.988	.938	.833	.884	.958	.958	.833
Low	High									
High	Low									
Low	Low	.992	.917	.917	.883	.779	.917	.925	.975	.954

data it appears that perceptions of occupational status symbols were independent of both status and occupational level interests. One exception to this generalization can be seen in the last column of Table 5, "Which of these men would you want to be like?" The lowest correlation, .392, was between the ratings of the two high status groups. On the other hand, the correlations between ratings of the two low status groups was .954. This difference indicates that interest scores were a factor in the occupational preferences of high status groups, but not in the preferences of low status boys.

Perhaps occupational preferences were in the nature of wish fulfillment for these low status boys. It is possible that the preferences of these boys will become more "reality" oriented as they grow older. It is also possible that if lower status boys had been asked to state second preferences, the preferences might have been more closely related to their interests. Small's study of the vocational choice of adolescents (20) suggests such a possibility. Unfortunately the data collected in this study do not permit such an analysis.

1. *Interest Scores and Ideas of Social Class Behavior*

Responses to the Guess-Who test were analyzed separately for the same status-interest groups described earlier. Rank-order correlations between the ratings by the various status-interest groups for each item on the Guess-Who test are shown in Table 10.

Although there was variation in the size of correlations between status-interest groups on certain traits, the over-all finding, as indicated by the mean z 's, is that the groups' ideas of social class behavior were remarkably similar. There was no consistent trend for either similar interest or status groups to have higher correlations. Apparently, both interest scores and social status were independent of these ideas of social class behavior.

Findings discussed to this point make it increasingly clear that boys from all socioeconomic backgrounds had similar ideas of social class as related to occupation. Deviations from group patterns or "norms" must be explained, therefore, on some basis other than social class. The only responses in which social class bias appeared to have any consistent influence were those with personal references, i.e., occupational choice and responses to items on the "occupational level" interest scale.

2. *Interests and Reputation of Peers*

During the interview the Guess-Who test was administered with reference to other fifth grade boys in a particular school. These sociometric ratings were correlated with occupational level scores. The results are shown in Table 11.

Although correlations between occupational level interests and some aspects of reputation were rather low, a number of them were statistically significant. Boys with high interests were perceived as boys who avoid fights. Boys with low interest scores were rated as untidy, fond of fighting, followers rather than leaders, poor students, not wanted as best friends, and as bullies.

The ratings were summed, e.g., "tidy and untidy," and the totals were correlated with interest scores. The correlations are shown in the last column of Table 11. Interests were positively correlated with "avoid fights-fights," "good student-poor student," and "friendly-unfriendly." These correlations, though statistically significant, were too low to be of practical use for predictive purposes.

3. *Summary*

Occupational level interest scores of these fifth grade boys were independent of their perceptions of class symbols and of their ideas of social class

TABLE 10
IDEAS OF SOCIAL CLASS BEHAVIOR AND OCCUPATIONAL LEVEL INTEREST SCORES. RANK-ORDER CORRELATIONS BETWEEN INTEREST-STATUS GROUPS

Question	Status(S)-Interest(I) Categories					
	High(S)-High(I) Low(S)-High(I)	High(S)-High(I) High(S)- Low(I)	High(S)-High(I) Low(S)- Low(I)	Low(S)-High(I) High(S)- Low(I)	Low(S)-High(I) Low(S)- Low(I)	High(S)- Low(I) Low(S)- Low(I)
Tidy	.981	.892	.929	.929	.888	.896
Untidy	1.00	.917	.967	.917	.967	.867
No Fight	.925	.929	.858	.883	.846	.975
Likes to Fight	.883	.929	.858	.983	.946	.929
Leader	.750	.721	.833	.913	.708	.621
Follower	.767	.767	.825	.825	.804	.808
Active	.696	.867	.804	.708	.625	.588
Sedentary	.767	.038	.482	.050	.400	.488
Good Student	.892	.958	.858	.954	.862	.921
Poor Student	.946	.858	.908	.912	.954	.917
Best Friend	.892	.833	.856	.967	.804	.838
Not Best Friend	.442	.546	.300	.929	.817	.908
Friendly	.683	.892	.817	.700	.758	.950
Unfriendly	.333	.504	.504	.588	.696	.167
Like to Visit	.808	.929	.975	.921	.825	.958
Don't Like to Visit	.875	.642	.796	.913	.883	.738
Protect the Weak	.883	.821	.917	.929	.783	.642
Bully	.954	.983	.854	.929	.929	.846
Mean z	1.43	1.25	1.25	1.43	1.23	1.25

behavior. Interest scores appeared to be a factor only in the occupational preferences of high status boys. Reputation of peers, however, seemed to reflect, in part, the occupational level scores. Apparently the values of boys, in terms of likes and dislikes, were expressed in the boys' behavior. This behavior, in turn, was related to their reputation among their peers.

TABLE 11
CORRELATIONS BETWEEN OCCUPATIONAL LEVEL INTEREST SCORES AND REPUTATION OF
PEERS: N 243

Tidy	.089	Untidy	— .287**	Total	.089
Avoids Fights	.136*	Fights	— .208**	Total	.132*
Leader	.043	Follower	— .167**	Total	.017
Active	— .125*	Sedentary	.101	Total	— .063
Good Student	.107	Poor Student	— .142*	Total	.121
Best Friend	.029	Poor Friend	— .182**	Total	.051
Friendly	.089	Unfriendly	— .073	Total	.151*
Protects Weak	.042	Bully	— .130*	Total	.057
Like to Visit	— .004	Not Visit	— .034	Total	— .021

* Significant at .05 level.

** Significant at .01 level.

E. SUMMARY AND DISCUSSION

1. Problem

The purpose of this study was to investigate the relationship between socioeconomic status and children's occupational attitudes. Specific problems involved determining the relationship of socioeconomic status to the perception of symbols of occupational status, to ideas of social class behavior, and to reputation among peers; and the relationship of children's occupational level interests to the above variables.

2. Procedure

The Dreese-Mooney Interest Inventory for Elementary Grades was administered to a sample of 243 fifth grade boys from elementary schools within commuting distance of the San Francisco Bay Area. Next, a 50-minute interview was held with each boy. In the interview, the boys were presented a set of drawings depicting nine occupations from three different socioeconomic levels to which they were asked to relate certain social class symbols. Each boy was then twice given a Guess-Who test, once with reference to the occupational pictures and again in reference to his peers.

From the responses to the pictures and from the Guess-Who test, data were obtained concerning children's perceptions of class symbols and their ideas about social class behavior. Information regarding reputation of peers was also obtained from the Guess-Who test. Occupational level interest

scores were obtained by means of a scale derived from the Dreese-Mooney Inventory. Father's occupation, classified according to a revision of Roe's occupational scale, was used as the criterion of socioeconomic class.

3. Findings

a. Perception of class symbols. This group of fifth grade boys perceived certain social class symbols as being associated with occupations and these perceptions were fairly accurate (i.e., were similar to those observed in previous studies of socioeconomic status). Perceptions of two occupations, however, were consistently at variance with expectation. "Policeman" was ranked very high on most status characteristics relative to the other occupations; "salesman," very low.

With one exception, perception of class symbols appeared to be independent of the subject's own status. Boys from lower status homes tended more often to perceive the unskilled workmen as living in the nicest house than did other subjects. This relationship, however, was not clear-cut ($.01 < P < .05$).

b. Ideas of social class behavior. The findings indicate that the subjects had fairly well-defined expectations of how a boy from a particular socioeconomic background would behave. In general, these expectations were independent of the socioeconomic background of the subjects. The following exceptions were noted:

1. Subjects from Roe's Classes I, II, and III were more likely to perceive sons of upper status occupations as leaders than were subjects from the other levels.

2. A larger percentage of subjects from Class III indicated they would like to visit lower status homes than did subjects from the other levels.

3. More subjects from Classes I and II perceived boys from upper status homes as being concerned with the welfare of the underdog than did subjects from other classes.

4. Subjects from Class III were more likely to see boys from lower status homes as being protectors of the weak than were subjects from Classes I, II, and IV.

5. Subjects from Classes I, II, and IV were more likely to see boys from middle status homes as being followers or non-leaders than were subjects from other classes.

6. Subjects from Classes I and II were more likely to reject boys from lower status homes as friends than were subjects from other classes.

7. Subjects from Classes I, II, and IV tended more often to perceive boys from lower status homes as being unfriendly than did other subjects.

c. Socioeconomic status and reputation. Sociometric ratings of peers were independent of the socioeconomic status of the raters and of the children being rated. Boys would choose friends or be chosen by others as friends, for example, without regard to social class.

d. Occupational level interests and other variables. Scores on the occupational level interest scale were independent of the perception of class symbols and of ideas about class behavior. Interest scores were related to occupational preferences of upper status boys and not to those of boys from lower status homes.

There were rather low but significant correlations between interest scores and several aspects of reputation. Boys with interests similar to sons of professional and business men were regarded as boys who avoid fights. Boys with interest scores similar to those of sons of semiskilled and unskilled workmen were regarded as fond of fighting, followers, poor students, undesirable as best friends, and as bullies.

4. Discussion

The findings indicate that social class was a factor in the occupational attitudes of this group of fifth grade boys, although the relationship was more complex than has been concluded from studies such as Stendler's (21). These boys had definite ideas about how social class symbols are related to occupations. Their ideas, however, were independent of their own socioeconomic status. They had definite ideas about how children from different social classes behave, but their evaluations of the behavior of their peers were not related to their own social class or to the social class of their peers. With few exceptions, the perceptions of boys from lower status homes were remarkably similar to those of upper status boys.

How can the similarity of perceptions be explained? Do they reflect elementary curricula which hew to middle class values by stressing the desirable aspects of aspiring to higher level occupations and, by implication at least, deprecating lower level occupations? What are the effects of mass media such as television, which tends to glamorize professions? Does the similarity reflect a diminishing influence on the part of the father in the development of the values of the son? Certainly there is speculation that boys today are influenced less by their fathers' occupations than were boys of a generation ago. According to Erikson (5), the mother is currently the

most influential member of the family constellation in transmitting values to her son.

Although the preponderance of the findings presented in this report indicated similarity among various socioeconomic groups, findings reported elsewhere (in press) showed a clear-cut relationship between socioeconomic status and responses of the subjects on a children's interest inventory. The patterning of responses of the criterion groups used in developing the occupational level interest scale were markedly different for the high and low status groups. Lower status subjects tended to respond with *like* or *don't know*; higher status, with *indifferent* or *dislike*. Differences in the response patterning for the status groups suggest that responses of lower status children to personal reference items are in the nature of wish fulfillment, (i.e., they would like many things which they cannot afford). The fact that children's interests are related to occupational preferences of upper status boys and not to those of boys from lower status homes would seem to support such an interpretation.

There were two different kinds of responses given by the subjects. If a subject was asked to evaluate a class symbol in an abstract and impersonal manner, his own social status did not enter into his response. Perhaps his response was guided by what he perceived to be socially correct or acceptable. On the other hand, when the subject was asked to make personal decisions about his own choices and preferences, his responses were related to his socioeconomic background. Thus, his own values were involved in the latter instance, and these values were determined in part by his social status.

While reputation of peers was independent of the socioeconomic status of both children who were judges and children who were being judged, occupational level interest scores were significantly related to several aspects of reputation, particularly to those aspects which tended to be considered characteristic of boys from lower status backgrounds. Peers were evaluated in terms of whether their values and behavior corresponded to ideas of lower class behavior—not whether the peers actually belonged to the lower social class. These findings suggest that in order to understand the importance of social class values in the behavior of boys in this age group, knowledge of socioeconomic background alone is not sufficient. These fifth grade boys judged their peers on the basis of their values as they were reflected in their behavior, not on the basis of their social class.

Findings of this study are only partly in agreement with earlier studies such as those reported by Stendler (21) and by Newgarten (18). According to Stendler, socioeconomic status was closely related to awareness of class

symbols and to sociometric ratings of children's behavior, especially in the upper elementary grades. Newgarten found that children in Grades 5 and 6 tended to select as friends children of higher status for the first choice; for the second choice, children of their own status. There was no evidence in this study to support either of the above conclusions.

How can these discrepancies be explained? Have the children changed in the nine years since Stendler's study was published, or in the 12 years since Newgarten's report appeared, or is the difference due to variations in sampling and in analyzing data? Whatever the reason, the findings of this study indicate that boys of this age level do not conform to social class expectations as much as has been implied by previous reports. A boy may be aware of class expectations and still behave independently of them. Perhaps this is an encouraging trend.

Whether these findings represent a genuine change in the tendency to conform or whether they reflect differences in techniques of examining the data, the pessimism expressed in other studies (3, 8, 9, 14) over the pervasive influence of social class on children's behavior may no longer be justified. If there is a cultural change, the nature of future research becomes clear—to identify and to describe the factors which are producing these changes.

On the other hand, future research may reveal that social class becomes more important in children's attitudes and behavior as they grow older. There is evidence that this is a possibility (21). Then a major task, if society is to realize effective utilization of its human resources, is to discover ways of preventing class boundaries from becoming fixed. As Warner has suggested (31:297), the survival of our system of open classes depends upon a definition of the individual as a person who has choice about what he does.

REFERENCES

1. BORDIN, E. S. A theory of vocational interests as dynamic phenomena. *Educ. Psychol. Meas.*, 1943, **3**, 49-65.
2. DARLEY, J. G., & HAGENAH, T. Vocational Interest Measurement. Minneapolis: Univ. Minnesota Press, 1955.
3. DAVIS, A. Socialization and adolescent personality. In *Adolescence*. Forty-third Yr. Bk. of National Society for the Study of Education. 1944.
4. DREESE, M., & MOONEY, E. Interest Inventory for Elementary Grades. Washington, D. C. George Washington Univ., 1941.
5. ERIKSON, E. H. *Childhood and Society*. New York: Norton, 1950.
6. GINZBERG, E. Sex and class behavior. In *About the Kinsey Report*. (Geddes and Curie, Eds.) New York: New American Library, 1948.
7. ———. Conservation of Human Resources: Report. New York: Graduate School of Business, Columbia Univ., 1955.

8. HAVIGHURST, R. J., & TABA, H. *Adolescent Character and Personality*. New York: Wiley, 1949.
9. HOLLINGSHEAD, A. B. *Elmtown's Youth*. New York: Wiley, 1949.
10. HOROWITZ, E. I. The development of attitudes toward the Negro. *Arch. of Psychol.*, 1936, **194**, 29-31.
11. JENNINGS, H. H. *Leadership and Isolation*. New York: Longman Green, 1947.
12. JERSILD, A. T., & TASCH, R. J. *Children's Interests*. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949.
13. LEHMAN, H. C., & WILTY, P. A. *The Psychology of Play Activities*. New York: Barnes, 1927.
14. LYND, R. S., & LYND, H. M. *Middletown in Transition*. New York: Harcourt Brace, 1937.
15. MACDONALD, M., MCGUIRE, C., & HARRINGBUSH, R. J. Leisure activities and the socio-economic status of children. *Amer. J. Sociol.*, 1949, **54**, 505-519.
16. MILES, L. T. *Children's preference for school and out of school activities*. Unpublished doctoral dissertation. Univ. California, 19-.
17. MOSIER, H. P., DUBLIN, W., & SHELSKY, I. M. A proposed modification of the Roe Occupational Classification. *J. Counsel. Psychol.*, 1956, **3**, 27-31.
18. NEWGARTEN, B. Study of children's friendships. *Amer. J. Sociol.*, 1946, **51**, 306-12.
19. RADKE, M., TRAVER, H. G., & DAVIS, H. Social perceptions and attitudes of children. *Genet. Psychol. Monog.*, 1948, **40**, 327-447.
20. SMALL, L. Personality determinants of vocational choice. *Psychol. Monog.*, 1953, **67**, No. 1 (Whole No. 351).
21. STENDLER, C. B. *Children of Brasstown*. Urbana: Bureau of Research and Service, Univ. Illinois, 1949.
22. STRONG, E. J., JR. *Vocational Interests of Men and Women*. Stanford: Stanford Univ. Press, 1943.
23. ———. *Vocational Interests 18 Years after College*. Minneapolis: Univ. Minnesota Press, 1955.
24. SUPER, D. E. *Appraising Vocational Fitness*. New York: Harper, 1949.
25. TRYON, C. M. Evaluations of adolescent personality by adolescents. *Monog. Soc. Res. Child Devel.*, 1939, IV (Whole No. 4).
26. TYLER, L. E. Relationships of interests to abilities and reputation among first grade children. *Educ. Psychol. Meas.*, 1951, **11**, 255-64.
27. ———. Development of "vocational interests": The organization of likes and dislikes in ten-year-old children. *J. Genet. Psychol.*, 1955, **86**, 33-44.
28. UNITED STATES EMPLOYMENT SERVICE. *Dictionary of Occupational Titles*. Washington, D. C.: Government Printing Office, 1939.
29. WALKER, H. M., & LEV, J. *Statistical Inference*. New York: Holt, 1953.
30. WARNER, M. L., MEEKER, M., & EELLS, K. *Social Class in America*. Chicago: Science Research Associates, 1949.
31. WARNER, W. L., et al. *Democracy in Jonesville*. New York: Harper, 1949.
32. WOFLE, D. *America's Resources of Specialized Talent*. New York: Harper, 1954.

School of Education
University of California
Berkeley 4, California

MODULATION OF SPEED OF RESPONSE WITH AGE*¹

Section on Aging, National Institute of Mental Health

JACK BOTWINICK, JOSEPH F. BRINLEY, AND JOSEPH S. ROBBIN²

A. PROBLEM

In a previous study an older and a younger group were compared for their decrement in rate of performance due to continuous task repetition. It was found that the younger group manifested a larger decrement, i.e., the rate of performance of the older group was more constant than that of the younger group (9). In this context Birren proposed the concept that with age there is a decrease in the adjustive range or "modulation of speed of response" (1).

There is now a growing body of information that is pertinent to this issue. It is known that: (a) elderly adults are slower in psychomotor responses than younger adults (e.g., 4, 13); (b) the slowing with age does not appear to be related primarily to motivation (8); (c) the slowness of older adults does not appear to be related primarily to peripheral neuromuscular events (4, 15); (d) the slowing with age may be related to a state of preparedness to respond (6, 15); (e) the age difference in speed of response is related to the difficulty of the task (e.g., 2, 3, 5, 7, 10); and (f) as already mentioned, younger adults decelerate response speeds with continuous repetition more rapidly than do older adults (9).

A question that needs to be answered in order to give a more complete picture of age differences in the range or modulation of speed of response is whether old adults can be as adaptively slow or slower than younger adults. The problem of this study, therefore, was to compare old and younger adults in their response to instructions for slow performance.

B. METHOD

1. Subjects

There were two age groups of male volunteer Ss. The older group comprised 29 Ss aged 65-81 years with a median of 72 years. The younger

* Received in the Editorial Office on May 6, 1958.

¹ Laboratory of Psychology, National Institute of Mental Health, National Institutes of Health, Public Health Service, U. S. Department of Health, Education & Welfare, Bethesda 14, Maryland.

² Grateful acknowledgement is due Dr. James E. Birren for suggestions relating to this study.

group comprised 34 *Ss* aged 18-32 years with a median of 22 years. Two of the older *Ss* resided in a home for the aged and the remaining subjects were community residents.

Nine of the older *Ss* and one younger *S* were foreign born. Education level for these *Ss* was at or equally distributed on either side of the median education for the age group. Both the median and mean education were approximately 11 years for the older group, and 14 years for the younger group. The largest age difference in education, perhaps, was the range; 12-20 years for the younger group and 4-19 years for the older group.

2. Procedure

Each *S* was instructed to write with a pencil as slowly as possible and without stopping, the phrase "New Jersey Chamber of Commerce." This phrase was chosen for the slow writing task because of its previous use in the reported literature (e.g., 12). Its merit is that no letter "i" or "t" is involved that needs dotting or crossing. As suggested in the literature, parallel lines were drawn one-half inch apart, and instructions were given to *S* to write the phrase within the boundaries as slowly as possible.

In addition to this, the procedure of the current study was expanded to measure usual or normal writing time of the phrase and also time of writing it as quickly as possible. The order of speed instructions of writing time that were administered was: (a) usual, (b) slow, (c) fast, followed by (d) a retest of slow writing.

No upper time limit was used for the slow writing task with the early *Ss*. Later, however, a time limit was imposed because of practical considerations including the comfort of *S*. The upper time limit was 20 minutes per phrase for both test and retest of the slow procedure, later reduced to 10 minutes. The score was simply the time, in sec., taken to write the phrase. The top score possible was 600 sec., even for those early *Ss* who took a longer time. A stop watch was used to time writing speeds to the nearest 1.0 sec. It is to be noted that this procedure is somewhat different from that used in previous studies (e.g., 12).

C. RESULTS

1. Reliability

The reliability of the measure of slow writing time was determined by the test-retest correlation of the first and second writings. For the older group the Pearson product moment correlation was .92. For the younger group, $r = .94$, and for both groups combined, $r = .95$. The coefficients

involving the younger group may be spuriously high because of six Ss who met the limit of 600 sec. imposed by the procedure. None of the older Ss met the limit. The reliability of the procedure, however, appears to be very good.

2. Age Differences

For each S the mean time score of the first and second writings under slow speed instructions, along with the time scores of normal and fast speed instructions, are presented in Figure 1. These scores were plotted on a log

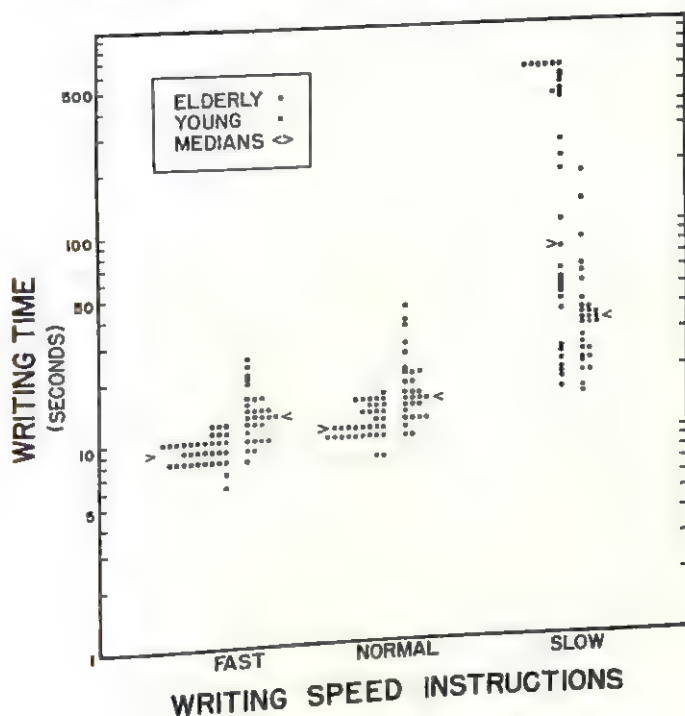


FIGURE 1

AGE DIFFERENCES IN WRITING TIME

Instructions were to write the phrase "New Jersey Chamber of Commerce" as fast as possible, at a usual or normal speed, and as slowly as possible. Each score represents an individual subject. Slow writing scores were means of two trials, with a maximum cut-off point of 600 sec. Scores are presented on a log scale.

scale in order that the distributions might be compared without resorting to an excessively large ordinate. It may be seen that slow writing scores for the younger Ss appear to be bimodal or to fall into two distributions—those who take 422 or more sec., and those who do it in 257 or 224, or less sec.

Accordingly, a non-parametric statistic was used to test the significance of age differences. The Mann-Whitney "U" test (11) indicated differences between age groups to be significant beyond the .01 level of confidence. When a test was made with only the lower scores of younger Ss, i.e., younger Ss who made scores of 257 sec. or less, no significant age difference was indicated ($p > .05$).

Another test of the hypothesis was to compare age groups with respect to the difference between the mean slow writing time and the normal writing time for each subject. When this was done for all the data, again the age difference was significant, $p < .01$. These difference scores, however, were distributed in a manner similar to that of the scores of slow writing and it was not possible to tell exactly in which end of the distribution some middle value scores fell. When "U" tests were performed with the difference scores of all the 29 elderly Ss and only with the smallest difference scores made by 19 younger Ss, and with the smallest difference scores made by 27 elderly and 17 younger Ss, age differences were significant at less than the .05 confidence level. When the test was made with the scores of all the elderly but only 18 of the younger group, $p > .05$. This latter test, however, was not considered as meaningful as the former tests because the cut-off was not as rational. It is concluded therefore, that these tests on the difference scores indicate significant age differences.

When both the data of normal and fast speed instructions in Figure 1 were compared between the two age groups, standard t -tests and "U" tests indicated that the younger group was significantly faster than the older group ($p < .01$). This suggested that not only is there with age a decrease in ability to do things quickly when required to, but there is also a slowing down of normal speeds in tasks that are as well practiced as a writing task.

3. *Rôle of Education*

Although the two age groups are comparable with respect to several aspects of amount of formal education, the range of years of schooling was less for the older group. To determine the rôle that education might play in the ability to write slowly or to inhibit motor speed, in each age group, years of education was correlated with the mean time scores of the first and second writings under the slow speed instructions. It was found that for the older group, $r = .39$ and for the younger group, $r = .31$. Only the correlation for the older group was statistically higher than zero ($p < .05$). The low correlations suggest that education is not a variable of special consequence in these data.

D. DISCUSSION

The data of this study are compatible with the generalization that older adults operate within narrower ranges of response speed than do younger adults. Compared with young adults, older subjects are neither as fast nor as slow, there is a constriction in the range of possible response speeds. Related to this generalization is the one that older adults do not inhibit as well as do younger adults. The present task was considered in several previous studies (e.g., 12) to be an index of motor inhibition.

If the slow writing data of young and elderly Ss in Figure 1 are examined with respect to the combined distribution of scores, it may be seen that it is possible to order the data in a four-category classification. Time scores of 15 to 25 sec. (Ranks 1 through 16) include both elderly and younger Ss. These Ss appear to be either unable or unwilling to inhibit the tendency to write quickly. Time scores over 25 sec. to approximately 38 sec. (Ranks 17 through 29) include older Ss primarily, and perhaps indicate an effort to inhibit but with relatively little success. Time scores over 38 sec. to approximately 180 sec. (Ranks 30 through 47) again include both elderly and younger Ss and suggest a more successful, but still limited, attempt in inhibition. Time scores over 180 sec. (Ranks 48 through 63) included only younger Ss and these may be considered to inhibit successfully.

Accepting the generalization that older adults operate within narrower ranges of speed of response than do younger adults, it may be asked whether this is more reasonably explained by an altered volitional or motivational preference, or by a changed physiological condition. With respect to functions requiring fast response, it was indicated in a previous study that an altered motivational state may not be an appropriate explanation of the slow-down with age (8). With respect to functions requiring slow response, no data are available bearing directly on its relation to motivational states. Fortunately, however, a test of level of aspiration was administered to the elderly Ss in this study.³ As a tentative answer to the question, level of aspiration was assumed to be similar or related to motivation, and the test score was correlated with the score of slow writing speed. The correlation coefficient was not statistically larger than zero, suggesting that elderly adults may operate within a narrower range of response speeds due to factors other than motivational ones.

A question that is related, perhaps, to volition or motivation is one that

³ We thank Dr. Margaret Thaler, Walter Reed Army Institute of Research, Washington, D. C., for making these data available to us.

has to do with what was communicated to the Ss by the instructions of this experiment. Is it possible that the task had a different meaning for older Ss than for younger ones, or that the two age groups had different models? The question is essentially one of whether the older group *did not* or whether it *could not* inhibit. This has to do with the general question of control. Further studies might well make use of different instructional sets by demonstration, goal setting, and by use of varied speeds that are to be controlled.

There is the *possibility* in the present study that the younger Ss may belong to one of two populations, one of which may not differ significantly from the older group. If, in fact, there are two populations of younger adults, then we have no data or explanations to offer as to why this is the case. The generalization, however, regarding a narrow range of response speeds for the elderly would still be tenable because the younger group manifested significantly larger scores of the difference between slow and normal writing, even when only the lower part of the distribution, or lower distribution of scores of the younger group was considered.

If it had not been for the artifactual cut-off at 600 sec., it is possible that the data of the younger group would have been markedly skewed in one distribution rather than bimodal, or of two distributions. Meltzoff's⁴ data, collected and scored somewhat differently from those of the present study, indicated one markedly skewed distribution. His studies included a distribution of scores made by 80 university students. When this distribution is compared to that of the younger group in the present study, a similarity was found in the percentage of Ss who had low time scores. Sixty-five per cent in both groups did not inhibit beyond 300 sec., and 52 per cent of the Ss in the current study and 48 per cent in Meltzoff's data did not inhibit beyond 180 sec.

It is possible that a narrowed range of response speeds within which older adults may modulate or adjust their responses has safety or protective features associated with it. Singleton (14) has hypothesized that older persons, as compared with younger ones, are not as efficient in anticipatory planning of a series of movements. The reduced speed of fast responses may minimize errors in serial planning during a complex task. It is possible that the increased speed of slow responses may also minimize errors. With the hypothesized age changes in associative fluency (1), and "capacity to form and hold the mental model for action" (14, p. 231), there may be a need in the elderly for a minimum response rate, below which, continuity of serial sequences are

⁴ We thank Dr. Julian Meltzoff, Chief Clinical Psychologist, V.A. Outpatient Clinic, Brooklyn, New York, for making these data available to us.

jeopardized. A limitation in changes of speed, i.e., acceleration and deceleration of response speeds, may minimize errors in the elderly associated with monitoring corresponding changes in the perceptual environment. The question remains, however, whether these behavioral changes with age are adaptive due primarily to factors of experience or are a necessary consequence of some yet to be defined change in neural structure.

E. SUMMARY

The question was raised as to whether older adults can be as adaptively slow or slower than younger adults when instructions required slow performance. It was hypothesized that older adults might not be as slow because they tend to operate within narrower ranges of response speed.

Each *S* was required to write in normal fashion, the phrase "New Jersey Chamber of Commerce." He was then required to write it as slowly as possible, then as fast as possible, and again as slowly as possible. *Ss* were 63 male volunteers divided into two age groups; one group aged 18 to 32 years ($N = 34$) and the older one, 65 to 81 years ($N = 29$). The median ages were 22 and 72 years respectively.

The results were that younger *Ss* were both faster and slower than older *Ss* as required by the instructions. The data are discussed in the context of their possible relations to planning, associative, and monitoring behavior.

REFERENCES

1. BIRREN, J. E. Age changes in speed of simple responses and perception and their significance for complex behavior. In *Old Age in the Modern World*. Edinburgh & London: Livingstone, 1955. (Pp. 235-247.)
2. BIRREN, J. E., ALLEN, W. R., & LANDAU, H. G. The relation of problem length in simple addition to time required, probability of success, and age. *J. Geront.*, 1954, **9**, 150-161.
3. BIRREN, J. E., & BOTWINICK, J. Rate of addition as a function of difficulty and age. *Psychometrika*, 1951, **16**, 219-232.
4. ———. Age differences in finger, jaw, and foot reaction time to auditory stimuli. *J. Geront.*, 1955, **10**, 429-432.
5. ———. Speed of response as a function of perceptual difficulty and age. *J. Geront.*, 1955, **10**, 433-436.
6. BOTWINICK, J., BRINLEY, J. F., & BIRREN, J. E. Set in relation to age. *J. Geront.*, 1957, **12**, 300-305.
7. BOTWINICK, J., BRINLEY, J. F., & ROBBIN, J. S. The interaction effects of perceptual difficulty and stimulus exposure time on age differences in speed and accuracy of response. *Gerontologia*, 1958, **2**, 1-10.
8. ———. Effects of electric shock motivation on reaction time and its relation to age. *Amer. J. Psychol.*, 1958, **71**, 408-411.
9. BOTWINICK, J., & SHOCK, N. W. Age differences in performance decrement with continuous work. *J. Geront.*, 1952, **7**, 41-46.

10. CLAY, H. H. The relationship between time, accuracy, and age on similar tasks of varying complexity. *Gerontologia*, 1957, **1**, 41-49.
11. MANN, H. B., & WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statis.*, 1947, **18**, 50-60.
12. MELTZOFF, J., & LEVINE, M. The relationship between motor and cognitive inhibition. *J. Consult. Psychol.*, 1954, **18**, 355-358.
13. MILES, W. R. Psychological aspects of ageing. In E. V. Cowdry (Ed.), *Problems of Aging*. Baltimore, Md.: Wilkins & Wilkins, 1942. (Pp. 756-784.)
14. SINGLETON, W. T. Age and performance timing on simple skills. In *Old Age in the Modern World*. Edinburgh & London: Livingstone, 1955. (Pp. 221-231.)
15. WEISS, A. D. The locus of reaction time change with set, motivation, and age. *J. Comp. & Physiol. Psychol.*, in press.

Section on Aging

Laboratory of Psychology

National Institute of Mental Health

Bethesda 14, Maryland

THE RELATIONSHIP BETWEEN DELAYED SOCIALIZATION AND TRAINABILITY IN GUIDE DOGS*

Guide Dogs for the Blind and Roscoe B. Jackson Memorial Laboratory

C. J. PFAFFENBERGER AND J. P. SCOTT¹

A. INTRODUCTION

The senior author since 1945 has been Chairman of the Puppy Testing Committee of a project for selecting puppies to be raised and trained as guide dogs for the blind. At the time this project was first set up barely 25 per cent of dogs trained could pass the final tests, and a great many failed because of "inability to take responsibility." This meant that the dogs were not able to exercise independent judgment in meeting unusual situations, even when related situations had been covered by the training procedure. Such dogs would react well in training, but lost confidence and became confused when required to lead a trainer who was actually blindfolded.

In an effort to remedy this situation a program of puppy testing and selection was set up in consultation with staff members of the Division of Behavior Studies at the Jackson Laboratory and Dr. C. S. Hall, then at Western Reserve University. This program was put into effect and careful records were kept of the performance of the animals.

When we analyzed these data we found that an experimental environmental variable had been accidentally introduced and that it was producing important effects on the results of training. The puppies which were kept in the kennels after the early testing program showed a much higher percentage of failures than those removed to homes immediately. This result has considerable bearing on the studies of the effects of early socialization. The results of the selection program will be discussed in another paper.

B. SUBJECTS

As shown in Table 1, the puppies used were predominantly from the German Shepherd breed. There was a considerable rate of attrition between puppyhood and the time of training at one year of age. Of 185 puppies which passed the puppy tests only 124, or 67 per cent, were given guide dog training.

* Received in the Editorial Office on May 7, 1958.

¹ Data for this work was collected at Guide Dogs for the Blind, Inc., San Rafael, California, and analyzed while the senior author was a Guggenheim Fellow at the Roscoe B. Jackson Memorial Laboratory, Bar Harbor, Maine.

Those eliminated were divided into three approximately equal categories: died, unsuitable size, or retained for breeding. None of these factors involved are related to the early environment of the subjects. Of 64 puppies which failed the test, 30 were selected as controls to test the efficiency of the puppy testing program. The subjects thus comprise two groups, selected respectively for probable proficiency and lack of proficiency in the training program. The results for each population are analyzed separately.

TABLE 1
BREED OF ANIMALS GIVEN GUIDE DOG TRAINING

Breed	Passed tests		Failed tests		Total
	Became guide dogs	Failed	Became guide dogs	Failed	
German Shepherd	61	42	5	25	133
Labrador Retrievers	12	2			14
Boxers	5				5
German Short Hair	1	1			2
Total	79	45	5	25	154

C. METHODS

1. *Birth to Eight Weeks*

The standard type of training and environment can be briefly described as follows: from birth to five weeks of age the puppies are confined to the 24 x 40 inch whelping box and to the puppy room where they are born. There is an access door to the kennel run, but usually only the mother uses this as the threshold is 14 inches high. At about five weeks of age the puppies and mother are moved to the puppy kennel, where the exit is more convenient for the puppies to go in and out. They start playing in their runs, which have a concrete floor enclosed by a chainlink fence and are approximately 7 feet wide and 22 feet long. All kennels are radiant heated in winter and air conditioned in summer. At night the puppies are confined to the kennel, which is 8 feet long by 6½ feet wide. While the walls of each compartment or puppy room do not reach the ceiling, they are 4½ feet high and thus prevent the puppies from seeing anything that is going on outside of their own compartment until after they start playing in their runs. In the runs they see other puppies in adjoining runs and the mothers which still have unweaned litters. Thus they have an environment rich in dogs but meager in human contacts.

The kennel manager (female) cleans the kennel rooms daily, keeping the floor covered with torn newspaper for bedding. The runs are scrubbed down

each morning before the puppies are released from the compartment. The door from the run to the compartment is closed after the puppies are let out to play in runs, keeping puppies out while the compartment is cleaned. After that, puppies may enter and leave freely during the day. These runs are concealed from other human activities by walls on four sides, except for an access sidewalk.

The kennel manager feeds each bitch in the compartment with the puppies. As soon as the puppies are old enough to start trying to eat, food is put in for them, usually about a week or 10 days before they are weaned. Except for such medication as may be needed, their feeding, and the cleaning of the kennels, the puppies have no contact with people until they start playing in the runs, and then only incidentally as staff members or puppy testers pass their pens while going about their work on the access walk. This leads only to the runs and the incinerator, and so has a minimum amount of use.

At seven, nine, and eleven weeks of age the puppies receive their inoculations against hepatitis and distemper. During this procedure the puppy is handled and has more extensive contact than usual.

2. Eight to Twelve Weeks

Beginning on the first Thursday after the eighth week birthday the puppies are given the simple training and testing program summarized in Table 2.

TABLE 2
PUPPY TESTING PROGRAM*

Name of test		Age in weeks				
		8	9	10	11	12
Weighing, observe sense organs, teeth		x	x	x	x	x
Physical Measurement		x	x	x	x	x
Trainability:	come	x	x	x	x	x
	sit	x	x	x	x	x
	fetch	x	x	x	x	x
	heel					
Responses to novel or frightening situations through stimulation of:	ear	x	x	x	x	
	body	x	x	x	x	
	eye					x

* In addition pups were rated as to intelligent response, willing temperament.

This requires about 30 minutes each week, and during this time the puppy comes into contact with several testers, mostly women.

During the tests the puppies are taken out of the kennel runs for the first time. They are first taken to a weighing room, and then to a testing

room where they are given simplified dog training: learning to come when called, sit on command, walk on a leash, and to retrieve a ball.

The puppy is then placed in a small pen outside the kennel and tested for his reactions to sound and touch. On the final week the puppy is taken out into a simulated city block area and exposed to moving objects and obstacles similar to those it may meet later as a guide dog. This test most closely simulates its later experience as a guide dog.

Each puppy is rated by two observers on its performance, and in addition is given a general rating on "intelligent response" and "willing temperament."

In general, this experience briefly introduces the puppy to a new and much more complex social and physical environment. The puppy meets this experience alone, unsupported by its familiar social group, and outside its accustomed locality.

3. *Twelve to 52 Weeks*

Time of removal to private homes. This is the experimental variable. Puppies on schedule were promptly removed at 12 weeks. However, some were left in the kennels for periods varying from one to 11 weeks afterward, depending on whether homes were immediately available. As it happened, exactly 50 per cent of the puppies which passed the tests were kept in the kennels more than two weeks, while 73 per cent of the control group were kept longer (see Table 3).

TABLE 3
EFFECT OF HOLDING PUPPIES IN KENNEL BEYOND 12 WEEKS OF AGE

Weeks in kennel	Passed tests		Failed tests		Total	
	Became guide dogs	Failed	Became guide dogs	Failed	Became guide dogs	Failed
0-1	36	4	1	5	37	9
1-2	19	3	0	2	19	5
2-3	11	8	4	5	15	13
3 or more	13	30	0	13	13	43
Total	79	45	5	25	84	70

The puppies are individually raised in private homes by 4-H Club members. Instructions are given on how to care for the dogs and how to give them simple training. The home environments are not identical, but the dogs are sent out to them by random choice, and supervision is exercised so that only homes are used which provide good environment for the animals. In the home environment a puppy is isolated from other dogs, has close contact with several people but particularly with the one child who cares for it,

and has a rich physical environment closely related to the environment where it will eventually live.

4. *Fifty-Two Weeks*

The dog is returned to the kennels and is given an intensive training as guide dog over a period of 12 weeks. The animal may be discarded at any time during this period if it does not respond adequately to training. The dog is scored by the trainers on a series of tests. The reasons for failure are summarized in Table 4.

D. RESULTS

When the number of failures in training is analyzed according to the time at which the puppy is placed in the foster home (see Table 3 and Figure 1), it is seen that the record of dogs which passed the original tests and were placed immediately after 12 weeks is excellent, with approximately 90 per cent of success. The record of those placed in the second week is slightly poorer, but not significantly so, while those which have been retained in the kennel more than two weeks show a highly significant increase in the number of failures. A comparison of the puppies left in the kennels more than two weeks with those removed sooner, using the Chi-square test, gives a probability of obtaining such a result by accident of random sampling of less than .001.

The record of puppies which failed as puppies and were used as controls for the predictive value of the puppy tests is, as expected, much poorer in both early and late groups than the record of puppies which passed the tests. However, no statistical comparison is possible between early and late groups, since only eight out of 30 pups were removed from the kennel within two weeks after completion of the tests. Within this small sample there was actually a greater proportion of failures among puppies removed early. Adding all puppies together gives almost exactly the same Chi-square value for the effects of early and late removal from the kennels as was obtained with the "passed" group alone. It may be concluded that leaving the puppies in the kennels for more than two weeks has an important effect on their later trainability, at least in animals which pass the puppy tests.

Detailed causes of failure are shown in Table 4. Those animals which failed the puppy tests show a different distribution of reasons for failure as guide dogs, indicating that the tests have separated two genetically different populations. The tests evidently select out a large proportion of dogs which are timid or will not respond to training. On the other hand, refusal to

"take responsibility" appears to be primarily the result of being retained in the kennels.

Among puppies which passed the tests, the rate of failure for each cause is increased approximately five times as the result of retention in the kennel. It is suggestive that the last three items on the list, including nervous wet-

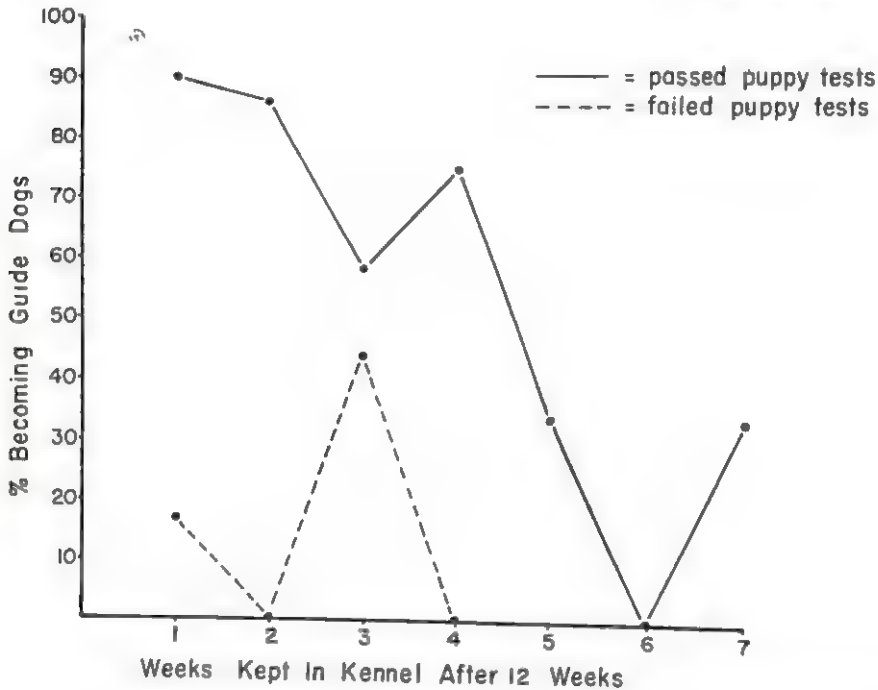


FIGURE 1

Puppies retained in the kennel more than two weeks after the conclusion of testing at 12 weeks of age show a markedly poorer performance as guide dogs. Puppies which failed the original tests are even poorer. Fluctuations in the curves are probably caused by small numbers.

ting, are found only in dogs retained in the kennels, but the numbers are too small to be significant.

E. DISCUSSION

1. Relation to Primary Socialization

In previous studies it has been shown that the puppy normally begins the process of primary socialization at approximately three weeks, and that if human contact is completely prevented the process becomes noticeably difficult by seven to eight weeks of age (4, 5, 6). By 12 weeks of age primary so-

cialization to human beings becomes so difficult that it is almost impossible to train a dog so that it is closely attached to human beings. This process of forming a close social relationship with human beings involves taking the puppy from the mother and litter during the period of socialization and making it entirely dependent upon the human handler.

TABLE 4
REASONS FOR FAILURE OF GUIDE DOG TRAINING

Reasons for Failure	Passed tests		Failed tests	
	Out before 2 weeks	Out after 2 weeks	Out before 2 weeks	Out after 2 weeks
Refuses to take responsibility	3	16		2
Will not train		3	4	7
Frightened of people, awnings, obstacles	2	11	3	7
Flighty, unstable	1	1		
Sharp (bites, aggressive)	1	3		1
Car sick		1		
Nervous wetting or defecation		3		
Too soft (over reacts to correction)				1
	7	38	7	18

The dogs in the present experiment were not entirely deprived of human contact during the first eight weeks, but neither were they given any experiences which would make them completely dependent on human handlers. Dogs reared in this fashion till later life tend to become increasingly timid toward human beings, particularly toward strangers. During this period they are usually strongly interested in the food brought by the caretaker, and show no excessive fear or attraction.

From eight to 12 weeks puppies are given an opportunity for closer contact with human beings. This means that the development of a deeper social relationship is begun only at the very end of the period of primary socialization. Furthermore, this training is carried on for a period of only one half hour per week. It is surprising that this amount of contact will produce a normal dog. Whether it would be sufficient if continued into adult life is not answered by the data.

Under the regular system, intensive socialization is delayed until 12 weeks of age, when the puppies are placed in homes. So far, the amount given appears to be sufficient for success in guide-dog training, at least for the selected group of puppies which pass the tests.

The puppy which is retained in the kennels two weeks longer suffers in two ways. The moderate degree of socialization experience involved in the testing procedure is broken off, and intensive socialization is delayed still farther beyond the end of the natural period. It is not clear whether age is the factor involved, or the breaking off of the early relationship. At any rate, the increase in the failures due to refusal to take responsibility indicates a disturbance of social relationships.

2. Effect of Poor and Rich Environment

The fact that many of the pups fail because of fearful reactions, in addition to the lack of confidence alluded to above, raises the possibility that the pups may not have had a sufficient variety of early experience in order to develop familiarity and confidence in new situations.

Hebb (2), Melzack and Scott (3), and Thompson and Heron (7) have offered data showing that animals reared in barren environments are unable to deal effectively with the problems of life in a rich adult environment. The environments in the experiments described by these authors are barren both in a physical and a social sense. By contrast, puppies reared as guide dogs have for the first eight weeks a barren physical environment, a relatively barren one regarding human contacts, and a rich one in respect to other dogs. From eight to 12 weeks they are taken out for brief weekly periods into an environment rich in physical contacts and human social contacts. The life in the foster home is similarly a rich one in these respects. If we assume that the poor results with animals kept in the kennels is due to a barren physical and human environment, we must conclude that there is a critical period during which an animal must make contact with a rich environment if it is to have a permanent effect. This explanation is closely related to the next.

3. Critical Periods of Learning

Thorpe (8) has found that in the life of song birds there is a critical period in which the song can be learned from the adults, and that this cannot be done earlier or later. Tinbergen (9) observed that Eskimo dogs in Greenland learned territorial boundaries within a week of the time of sexual maturity. This, however, occurs much later in life than the phenomenon found in guide dogs.

It is possible that there is a critical period at approximately 12 weeks for learning and adjusting to a new physical environment. This is suggested by some of our observations of puppies reared in large fields. These animals

tend to stay close to the nest box until approximately 12 weeks of age, and then begin to wander farther away.

4. *Break in Socialization*

The senior author of this paper is inclined to believe that the phenomenon produced here is the result of a break in the process of socialization. During the testing period the puppies are being given only a minimum amount of socialization each week and, if this is discontinued for two weeks or longer, the animal may forget its early training in this respect and be unable to relate it to later experience. With this in mind, puppies which now must be left in the kennels beyond 12 weeks are being given the regular tests each week, thus providing human contacts and socialization. The results should decide between the hypothesis of a break in socialization and those hypotheses which depend on maturation.

In conclusion, we can say that the experiment definitely demonstrates an effect of delay of normal early social experience on later behavior. It does not provide a clear-cut differentiation between various theoretical explanations which might be responsible for it.

The phenomenon itself is an interesting one, not only because of its practical use in the rearing of guide dogs, but because of its implications for human behavior. The "taking of responsibility" required of the guide dog is essentially similar to one kind of social relationship which is demanded of adult human beings. The guide dogs which fail for this reason tend to be (in human terms) somewhat socially immature and overdependent.

The dog's experience may be compared to that of a child in an orphanage, well cared for but having no special contacts with adults except for a group of visitors who come once per week and take him for a short excursion into the outside world. In one case they finally take him away permanently to a new home outside. In the other, they forget to come for several weeks, but finally come again and suddenly take him to a new home.

The studies of Bowlby (1) on the effects of deprivation of maternal care on children raised in normal homes indicate that serious emotional disturbances result. In the above case, where the relationship set up is a more superficial one, we would expect a less drastic effect. Puppies treated in this fashion appear to get along fairly well in their foster homes, and it is only when they face the strain and responsibility of guide-dog training (which involves adjusting to still another group of people and new circumstances) that they fail.

F. SUMMARY

1. Dogs trained as guide dogs for the blind frequently fail in training at one year of age. The most frequent cause of failure is failure to take responsibility for the blind person in situations requiring independent judgment.

2. This paper describes an experiment in which factors of heredity and early experience were varied.

3. Two hundred forty-nine puppies, predominantly of the German Shepherd breed, were given a series of aptitude tests at eight weeks and continuing through 12 weeks of age. Of these, 185 passed and 64 failed. One hundred twenty-four of the first group and 30 of the second were given guide dog training at one year of age.

4. Performance as a guide dog involves the development of a close social relationship with a human being. The development of this relationship occurs as follows: (a) 0-8 weeks—kennel rearing; casual contact with caretaker during cleaning and feeding; (b) 8-12 weeks—kennel rearing; one-half hour weekly periods of intensive socialization as puppies are tested; contact with several persons; (c) 12-52 weeks—home rearing; intensive socialization with a child and its family; (d) 52 weeks—return to kennel; intensive training in the guide dog relationship with a trainer; transferral of this relationship to the blind owner.

The development of this relationship may be compared to that of a child reared in an orphanage, removed to a foster home, and finally leaving home as an adult.

5. An environmental variable is introduced when dogs are left in the kennel beyond the age of 12 weeks. Since training is discontinued during this time two factors may affect these dogs: (a) age at which adjustment is made to the foster home, (b) a break in the developing relationship with human beings.

6. Puppies which passed the aptitude tests are significantly different from those which failed, both in the number of failures as guide dogs and the reasons for failure. This supports the validity of the original tests.

7. Of the group which passed, 50 per cent were kept in the kennels two weeks or more beyond the age of 12 weeks. Those kept longer than two weeks show a highly significant increase in failures as guide dogs, all causes of failure being multiplied about five times. Numbers in the "failed" group were not sufficiently large to make a statistical comparison, but these additions to the total do not change the result.

8. It is concluded that failure to develop a satisfactory relationship as a guide dog is produced both by the factor of heredity and that of early experience.

9. Two hypothetical explanations are offered for the effect of early experience: (a) a break in the development of a social relationship with consequent emotional disturbance, (b) a critical period for learning the adult social environment (or for adjusting from a "poor" to a "rich" social environment). Either one or both factors may be involved.

10. The result can be compared to the development of social relationships as experienced by children reared in orphanages.

REFERENCES

1. BOWLBY, J. *Maternal Care and Mental Health*. Geneva: World Health Organization, 1951.
2. HEBB, D. O. The effects of early experience on problem solving at maturity. *Amer. Psychol.*, 1947, **2**, 306-307.
3. MELZACK, R., & SCOTT, T. H. The effects of early experience on the response to pain. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 155-161.
4. SCOTT, J. P. The process of socialization in higher animals. In: *Interrelations between the Social Environment and Psychiatric Disorders*. New York: Milbank Memorial Fund, 1953.
5. ———. Critical periods in the development of social behavior in puppies. *Psychosom. Med.*, 1958, **20**, 42-54.
6. SCOTT, J. P., FREDERICSON, E., & FULLER, J. L. Experimental exploration of the critical period hypothesis. *Personality*, 1951, **1**, 162-183.
7. THOMPSON, W. R., & HERON, W. The effects of early restriction on activity in dogs. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 77-82.
8. THORPE, W. H. *Learning and Instinct in Animals*. London: Methuen, 1956.
9. TINBERGEN, N. *The Study of Instinct*. Oxford: Clarendon Press, 1951.

Guide Dogs for the Blind, Inc.
San Rafael, California

PERCEPTUAL FUNCTIONING IN THE DUCKLING*

Department of Psychology, Queens College, Flushing

NICHOLAS PASTORE

A. PROBLEM

Information on the visual discriminations of young animals born or hatched in a relatively mature state would be valuable in determining the degree to which perceptual functioning is influenced by autochthonous and learning factors. It would be generally conceded, in a first approach to the problem, that the closer to birth a perceptual datum is manifested the more likely that autochthonous factors are the principal causal agents. For example, if that autochthonous factors are the principal causal agents. For example, if size or brightness constancy can be demonstrated in an animal at age 20 hours after an exposure of two hours to the relevant stimuli, then strong evidence for the autochthonous point of view has been secured. On the other hand, a similar criterion for a demonstration of the rôle of learning remains equivocal since postnatal maturational processes may be confounding influences (8, Ch. 3). Moreover, subjecting young animals to light deprivation may impair the unfolding of maturational processes and directly influence perceptual development. Consequently, retardation in the manifestation of perceptual achievements in light-deprived animals is insufficient to demonstrate the rôle of learning. The contamination of maturational with learning factors is an error in methodology in much current work (Cf. 15). Investigation of the perceptual achievements of animals hatched in a relatively mature state affords another avenue to the problem.

The experiments to be reported here were motivated by the foregoing considerations. More specifically, confirmation of previously obtained results on the perceptual functioning of ducklings (14) and their extension to other perceptual problems were the objectives. The new problems involved the selection of a triangle on the basis of its overall shape¹ and brightness constancy.

B. SUBJECTS

Three incubator-hatched Pekin ducklings, acquired three days after hatching.² Training was begun immediately upon reception of the animals.

* Accepted for publication by Leonard Carmichael of the Editorial Board, and received in the Editorial Office on May 8, 1958.

¹ For the relevant avian literature on this problem see Ref. 17.

² I wish to thank Dr. H. E. Moltz, Brooklyn College, for the ducklings.

C. APPARATUS AND STIMULI

A stimulus platform (9" wide and 5" high) with three food wells positioned 16" away from a cage.

*a. Stimuli for the form problems.*³ (a) Triangles, rectangles, and trapezoids of various dimensions cut from tinplate (for examples, see Figures 1 (I *a, b, c*)). Each cut-out had a rectangular stub which was bent to an angle of 90 degrees. This stub provided support for the figure and also covered the food well. New cut-outs of the figures were introduced in critical runs in order to remove possible idiosyncratic cues. (b) Triangles, squares, rectangles, circles, and parallelograms of various dimensions cut from blue, green, white, and yellow cardboards (for examples, see Figures 1 (II *a, b, c, d*)). These solid figures were placed on tinplates or black cardboards so that they could be rotated to any position. The background for a figure measured $2\frac{1}{4}$ " by $2\frac{1}{4}$ ". An alternative parallelogram of area equal to that of Figure 1 (III *d*) but with a long diagonal 2" in length was substituted when necessary. Thus, the substitute figure could be set with its long diagonal in a vertical position without overlapping the tinplate background. Rotation of a figure was achieved by inserting a brad in the center and through a hole in the tinplate. A washer placed behind the tinplate made it possible to keep the figure flat against the front of the tinplate. The head of the brad, which was exposed, could not have been a distinctive differential cue. Moreover, it had the advantage of directing the duckling's aim to it. Nevertheless, as a check, the brad or axle was affixed behind new figures so that the head did not protrude at the front. The tinplate backgrounds were often varied among the positive and negative figures, thus eliminating possible background cues as determinants of the bird's response. When set on a black background, the figure was affixed to it by a drop of glue. Since the background could be rotated, four different orientations of the figure were possible. The figure would be detached and reglued for other orientations. (c) Outline figures of triangles, squares, rectangles, circles, and parallelograms of various sizes drawn in India ink on white or black cardboard backgrounds. The contours were continuous or discontinuous in varying degrees (Figure 1 III *d, e, f, g, h, i, j*). Moreover, in some figures (III *a, k*) there were changes in the background; in other figures (III *b, c*) there were reversals in figure-ground relationships; still other figures (III *l, m*) were solid or contained three continuous outline shapes. All figures in this series had their counterparts in squares, rectangles

³ The choice of figures was guided by discussions in Ref. 2, 3, 6, 9, 10, 16, 17.

and sometimes parallelograms. Figures II *a*—*k* and their counterparts were drawn so as to enclose equal areas. Each background card (a $2\frac{1}{4}$ " square), with a figure drawn on it, could be placed in one of four orientations. For instance, Figure III *d* could be placed with the apex pointing up, to the right, to the left, and down. Figures III *e*, *f*, *g*, *h*, *i*, *m* were also drawn so that various acute angles were formed between the base of each figure and a horizontal line. There were other figures in which their placement on the

FIGURES

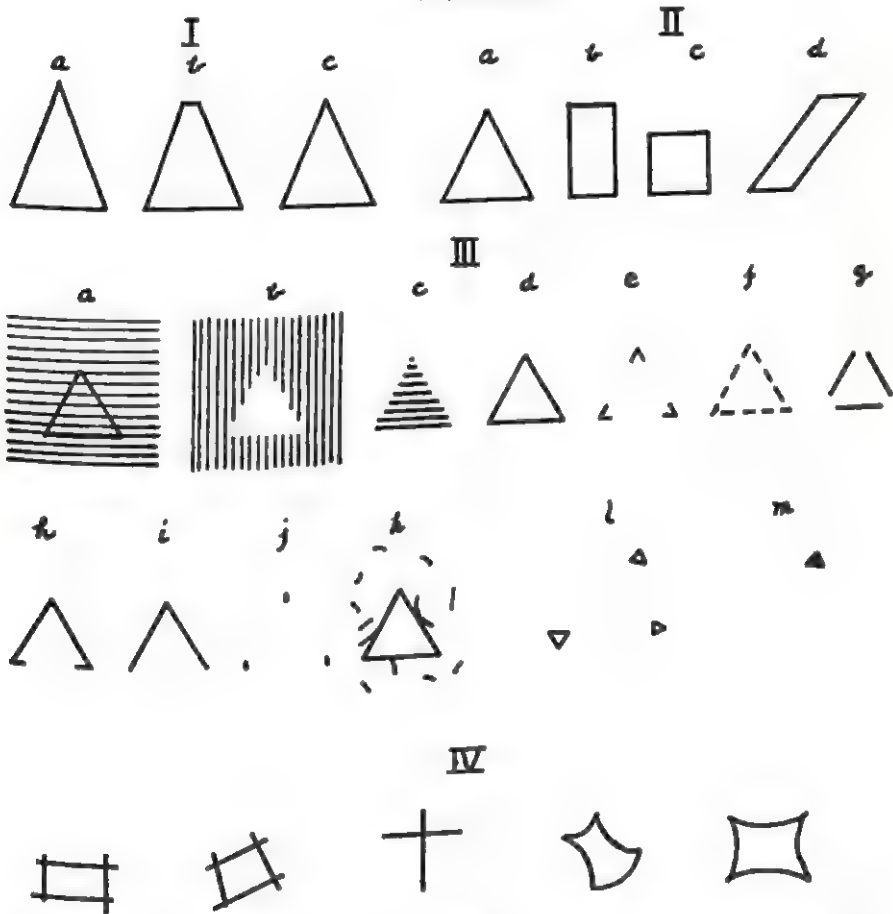


FIGURE 1

The actual altitude of Figure I *a* was 2"; this is the scale for all figures. Figures I were cut from tinplate, Figures II from cardboard. Figures III and IV were outlined in India ink (except for Figure III *m*, which was solid). The width of contour was 0.1". The width of striations in Figures III *a*, *b*, *c* was 0.05".

background was varied. For example, the triangle in Figure III *m* could be in the lower left hand corner in one trial, in the upper right hand corner in the next trial, and so on. (*d*) Various outline "nonsense" or other geometric figures (series IV figures).

b. Stimuli for the horizontal vs. vertical striations problem. A pair of $\frac{1}{2}$ " black stripes placed in a vertical or in a horizontal orientation on each of three square tinplate backgrounds.

c. Stimuli for brightness constancy. A pair of white and dark gray cards; a pair of cards of two shades of gray. The two shades of gray were No. 2 and No. 4 in the gray series manufactured by the Color Aid Co. (New York). They were matched respectively to shades No. 7 (light gray) and No. 5 (medium gray) in the Munsell gray series. The reflectance of Munsell No. 7 is 43.1 per cent and that of No. 5, 19.8 per cent. The cards, measuring $2\frac{1}{2}$ " square, were placed over the end food wells of the stimulus platform with a partition intervening. The length of the partition forced the bird to make a choice between the two cards when at least 4" away from the platform. The partition also made it possible to direct extra illumination to half of the platform, leaving the other half in shadow. The amount of incident light above the food well on the illuminated side was 42 footcandles and, for the half in shadow, two footcandles. With the No. 4 card in the illumination, the use of a reduction screen showed that it was considerably brighter to the *E* than the No. 2 card in the shadow.

Fresh stimuli were introduced frequently in all of the above problems, except those involving tinplate figures, in order to eliminate possible differential cues resulting from smudges produced by the bird's bill. Smudges on the tinplate figures were wiped off.

D. PROCEDURE

After learning to push aside a stimulus for food in the well below (a readily learned task), the bird was introduced to the form problems in which the triangle, always the positive figure, was contrasted to two negative stimuli. In the problem involving tinplate cut-outs the erect triangle (Figure I *a* or I *c*) was set against two rectangles or two trapezoids (Figure I *b*). After learning this discrimination the bird was introduced to the problems involving the choice of solid and outline triangles.⁴ The contrasting negative figures included various combinations of squares, rectangles, parallelograms, and

⁴ At the beginning of the series involving solid cardboard figures, the figures were unequal in area; figures equal in area were introduced subsequently. The specific changes are mentioned in the *Results* section.

circles. For example, a particular stimulus combination for one block of trials would consist of Figures II *a*, *b*, *c*. In the various stimulus combinations for these problems the triangle would be presented in an erect orientation first. Then, in subsequent trials, only the triangle was rotated through various positions (including the inverted position); in further trials the negative figures themselves were rotated and, finally, all figures were rotated. Of course the tinplate figures were never rotated.

The problem of discriminating horizontal striations from vertical striations and the brightness constancy problem were given to the birds directly following the successful selection of a triangle from a pair of trapezoids (tinplate cut-outs). The initial training in the brightness constancy problem involved a discrimination of a white card from a dark gray one under the condition of equal illumination. One bird was trained to respond to the white card as the positive stimulus, the other two to the dark gray card as the positive stimulus. After this brightness discrimination was learned, the birds were tested in the constancy set up. Subsequently the No. 4 and No. 2 gray cards, previously described, were introduced. The latter pair of cards was substituted in order to use grays of some standard series and to remove the possibility that the animal was responding to the "glare" of the white card or its absence from the gray card.

The average reward per trial was a mash of milk and bread weighing about $0.01 +$ oz. From the outset, the birds were deprived of food over night. During the day the birds received most of their daily rations while in a training situation. At night they were satiated, after which food was removed. This procedure made it possible to secure as many as 500 trials in a day's training with a single bird. This number was the maximum obtained for any bird. Since we were not interested in obtaining information on learning rates on a fixed schedule of a given number of trials per day we deviated from tradition in the training of the animals. The earliest manifestation of their perceptual functioning in an experimental setting was our goal. The question of the degree of interproblem transfer awaits further research in which a larger sample of animals necessarily must be used.

Although there was no punishment for errors *E* occasionally spanked the bird lightly if it dashed from the cage for the platform without apparently examining the stimuli. Such spanking was successful in slowing up the bird and was subsequently discontinued. Usually, after obtaining food the bird returned to the cage of its own accord. *E* blocked off the cage opening with oversize cardboard so that the animal was not visually exposed to the operations entailed in rearranging or replacing stimuli. When the bird was to

be released for a trial the cardboard was swung around, providing egress from the cage and at the same time obscuring the bird's vision of E. The choice of well for the positive stimulus was random except when a position response was to be broken. Since there were always three stimuli over the food wells in the *form problems* the probability of a correct response was $\frac{1}{3}$; the probability of a correct response in the constancy problem was $\frac{1}{2}$. One exception will be noted later. Fifteen or more hits in a block of 20 trials was the learning criterion. The transfer criterion was 9 or 10 hits in 10 trials. A correction procedure was used throughout. When not in a training situation the ducklings were kept in a kitchen tub which was 14" deep and 17" by 19" at the top.

E. RESULTS

1. *Discrimination of Triangle from Trapezoids (Tinplate Cut-Outs)*

Proceeding from stimulus combinations of triangles and rectangles to stimulus combinations of triangles and trapezoids, the average number of trials for the three birds to select the triangle successfully was 713 (criterion trials included); the range was 810 minus 580. For the final 100 trials of the series, involving only discriminations of triangles from trapezoids, the average number of hits for the three birds was 89. One bird achieved all hits. These results, obtained by the age of six days, confirm those previously reported (14). It should be observed that these results do not necessarily demonstrate a response to the forms per se; the presence or absence of the apex, for example, could have been the effective cue.

2. *In the Horizontal vs. Vertical Striations Problem the Vertical Striations Were Positive for Two Ducklings*

The average number of trials required to reach criterion for the three animals was 47 with a range of 80 minus 20.

3. *Brightness Constancy*

The initial discrimination of the dark gray card from the white one required, under the condition of equal illumination, zero trials to reach the criterion (white card positive) for one bird and 20 and 40 trials for the remaining two birds (dark gray positive). After the ducklings transferred their discriminatory response to the positive card in the constancy set up, the No. 4 and No. 2 gray cards were introduced. The two birds directly transferred their response to the No. 4 card, under the condition of equal

illumination, and to the same card in the constancy set up.⁵ For those trials in the constancy situation involving the pair of gray cards, the two birds together made only 9 errors in 80 trials. The results for the two ducklings concerned with the perception of triangular forms (cardboard and outline figures) are best presented individually.

a. Duckling 1. An equilateral triangle and two squares ($s = 1\frac{1}{8}$ " for the three figures) were cut from dark blue cardboard and mounted on tinfoil backgrounds. At first, the triangle was erect and the bases of the squares were horizontal. Twenty trials were required to reach the criterion. In the following trials, in blocks of 20 and then 10, the triangle was rotated so that the apex pointed to the left, down, up, and to the right. In these 60 trials, there were 55 hits. In the next series of trials, measurements of new figures were changed so that the area of the triangle was intermediate between that of two squares. Proceeding from trials in which the triangle was erect to trials in which it was rotated and, finally, to trials in which all figures were rotated the bird achieved 127 hits in 175 trials. In the next series of 250 trials, measurements of the triangle were varied, and circles were substituted for the squares in 110 of the trials. As before, all figures were rotated with the obvious exception of the circles. Measurements for a pair of the 11 different stimulus combinations were: a triangle with base 1" and height $1\frac{1}{4}$ ", two squares with sides of $\frac{3}{4}$ " and $1\frac{1}{4}$ "; a 1" equilateral triangle, two circles with diameters of $\frac{3}{4}$ " and $1\frac{3}{16}$ ". There were 234 hits in these 250 trials. Thus, from the bird's introduction to this problem to the end, there were only 83 errors in a total of 525 trials. Apparently there must have been considerable transfer from its previous successful discrimination of a triangle from trapezoids (tinfoil cut-outs). The figures in these trials were not of equal areas. *E* was interested in establishing the fact that the bird's response to the triangle did not depend on differences in the lengths of either the sides or the perimeters of the figures. A discrimination between a triangle and a square of the same area need not be based on form as such; the necessarily different lengths of the sides of the two figures could be the discriminative cue.

New figures of green cardboard, all of equal areas, were now introduced (Figures II *a, b, c*). Through successive trials the triangle was rotated from the erect to all positions and then all figures were rotated. There were 186 hits in the 225 trials of this series.

At this point, it is not clear whether the bird is responding to the triangle

⁵ One bird, the one for which white was the positive stimulus, met with a fatal accident at age 10 days.

as a whole or to some part of it. For instance, the triangle has an acute angle whereas the square and rectangle obviously do not.⁶ Accordingly, the rectangle was replaced by a parallelogram (Figure II *d*) the acute angle of which was equal to an acute angle of the triangle. The three figures were cut from bright yellow cardboard. In the 175 trials, trials in which all figures were rotated, there were 139 hits. Half the errors occurred in the initial trials—the animal appeared to be frightened by the introduction of the new color. When white figures (II *a, b, d*) on black background were substituted, the triangle was chosen in 52 of 60 trials. Finally, a blue square was contrasted to two blue circles, the areas of which were larger and smaller than that of the square. The fact that the square was chosen in 9 of 10 trials indicated more firmly that the bird's response to the triangle did not depend simply on the avoidance of the typical negative figures.

From the time of its introduction to solid figures on tinplate to the end the bird, by the age of 12 days, achieved 83 per cent hits in an over-all total of 995 trials. There is little doubt that the animal was responding to triangularity of form as such. Further confirmation is to be had in the results obtained from the use of outline figures.

The outline figures for the first problem were an erect triangle ($b = 1\frac{1}{2}$ ", $h = \frac{7}{8}$ "), and two squares ($s = \frac{7}{8}$ " for one and $s = 1\frac{1}{8}$ " for the other). Twenty trials were required to reach the criterion. For the next block of trials, the measurements of the figures were changed. The measurements of the basic triangle can be read from Figure III *d*. A triangle, whatever the changes in its contour or background, was associated with a corresponding square and a corresponding rectangle; all figures were equal in area. After the first 100 trials in this block, trials in which the triangle was always erect and contour continuous, its orientation was systematically varied in four ways—erect, apex to the left, inverted, apex to the right. Orientations of the square and rectangle were changed similarly. The various stimulus combinations proceeded from a triangle with a continuous contour (Figure III *d*) through figures with alterations in contour and background (Figures *c, e, f, g, a*). Figure III *a* was drawn with vertical striations as well. This made it possible to secure various orientations of triangle, square, and rectangle so that striations on the three figures were either only vertical or only horizontal. There were 356 hits for the 400 trials in this block. In subsequent trials, involving Figures III *b, j* the bird failed to show any transfer or learning.

In the next block of trials, the platform was lengthened so as to permit

⁶ An uncontrolled cue in previous work on this problem.

the presentation of five stimuli. All triangles previously used (with the exception of Figures III *b*, *j*), Figure III *l*, and scalene triangles were contrasted to four figures. The four contrasting figures included series IV figures, rectangles, parallelograms, trapezoids, and others as well. In these trials the orientation and the placement of the triangle on the stimulus card were varied. There were 324 hits in 370 trials.

Three additional results for this same bird are of interest. For the first run of trials the following kinds of triangles were used: cut-outs from tinplate and affixed to white backgrounds, outlined in pencil, and in solid pencil (for an example, see Figure III *m*). These triangles—which varied in area, symmetry, orientation, and placement on the background card—had their counterparts in rectangles and squares of the same or different areas. The smallest triangle was isosceles ($b = \frac{1}{8}"$, $h = \frac{3}{16}"$). In the 220 trials there were 188 hits. When the placement of the triangle was varied over the background card the duckling obviously directed its aim to it. For the second run of trials there were seven stimulus combinations. In order of presentation they were: (a) Figure III *d* and two of Figure III *i*, (b) Figure III *d* and two of Figure III *h*, (c) Figure III *h*, a rectangle, and a square, (d) Figure III *d*, Figure III *h*, and a rectangle, (e) repetition of *a*, (f) repetition of *d*, (g) repetition of *c*. All figures enclosed equal areas. The orientations of the continuous and discontinuous triangles were varied. There were 88 hits in a total of 100 trials. Apparently, for a given stimulus combination the duckling chose the figure which most closely resembled a continuous triangle. These results further confirm the interpretation that the animal was responding to the form of the triangle since Figure III *h* has as many acute angles as Figure III *d*. For the third run of trials an inverted triangle was contrasted to two erect triangles (3 of Figure III *d*). The inverted triangle was the positive figure because there was little doubt of the animal's preference for an erect one. This problem was chosen because of the apparent difficulty it has offered other lower animals (Cf. 5). The duckling required 10 trials to reach criterion.

From its introduction to outline figures to the end the bird achieved 85 per cent hits in a total of 1130 trials. The trials involving the inverted and erect triangles are excluded from this computation.

b. Duckling 2. This bird was introduced to the same solid figures as those of the first one at age 8 days. It required 60 trials to reach criterion for selecting an erect triangle. In the next 60 trials it selected the triangle successfully despite rotations to four different orientations. Subsequently, the bird's performance rapidly deteriorated to a chance level, developing

strong position responses. Although the discrimination of a triangle from trapezoids (tinplate cut-outs) was successfully reviewed its discrimination, nevertheless, remained at a chance level when re-introduced to the first problem of this series. *E* conjectured that this bird had reacted sensitively to spankings for it was reluctant to leave the cage and when it did so, it appeared to pay little attention to the stimuli. Consequently, the bird was given gentle treatment combined with frequent hand feedings. In the 11 days of such treatment the bird was tested on the initial problem. Although its performance improved, it functioned at a low level of efficiency. Due to *E*'s time limitations, the bird was introduced to the problem of selecting an outline triangle. The procedure now corresponded to that of the first duckling. In 620 trials spread over four days, this bird achieved 395 hits. After a vacation of 10 days, the bird was run through another sequence of 250 trials, achieving 204 hits. In an additional 40 trials this bird, like the first one, selected the triangle of Figure III *b* on a chance basis only.

Although birds have a poor sense of smell, if not altogether absent, it was decided to run control tests anyway. A triangle was contrasted to two circles, with food in the three wells. The bird achieved 10 hits in 10 trials. The fact that the first duckling usually directed its aim to the triangle, the placement of which was varied over the background, further demonstrates the ineffectiveness of possible olfactory cues. In *E*'s preliminary experiments (14) no food was placed in any well. Upon pushing aside the triangle the bird had to turn to the *E* who then supplied food by hand. The bird's selection of the triangle was not affected by such feeding.

F. DISCUSSION

The early appearance of some basic aspects of perceptual functioning in the duckling—aspects which include the perception of the form of a triangle, brightness constancy, size constancy (14), discrimination of horizontal from vertical striations—is strongly suggestive of their autochthonous determination. Logically, of course, empirical factors might have been influential during or prior to the onset of experimentation. Steps that remain in obtaining a definitive decision along this line depend on further experiments and a clarification of the meaning of such phrases as “learning to perceive” or “the influence of learning on perception” (12). Since the duckling is sufficiently active at age 20 hours, it may be possible to secure the same results earlier in its life.⁷ Moreover, the use of outline figures from the outset of training

⁷ Subsequent to the experiments of this report, brightness constancy and size constancy were demonstrated in one duckling at the ages of two and six days

may advance its learning achievements to an even earlier age. Certainly it would be important to assign a larger sample of animals to several subgroups so that each subgroup receives one basic problem only. An extension of this research to those mammals born relatively mature—for example, the ungulates—should yield further relevant information.

Among other considerations in clarifying basic terms, a distinction between the origin and modification of a percept is essential (13). The bulk of the experimental work pertaining to the issue of *nativism vs. empiricism* actually deals with the question of modification. Unfortunately, from the fact of modification of a percept no direct inference can be drawn about its origin. In our experimental work we are concerned, so to speak, with the initial perceptions of the duckling.⁸ Whether such perceptions, once they have occurred, can be modified or influenced by variable sequences of training, experience, or learning would be separate experimental considerations.

Although comparisons between species are risky, the duckling's superiority, where information is available, is apparent. For example, a simpler version of the discrimination of a triangle from closely related geometrical figures has been a difficult one for the rat (2). The following factors might underlie the duckling's superior performance. (a) The bird is primarily a visual animal. (b) *E* used massed trials, the time interval between successive trials was small. For example, in one run, 160 trials were obtained in one hour. (c) Knocking over a stimulus is immediately followed by reinforcement or nonreinforcement. Actually, the act of knocking over the positive stimulus and that of scooping up food from a shallow well were integrated by the duckling into a single response. The estimated time for such an integrated response certainly was less than half a second. For a rat in the Lashley jumping apparatus the time interval between striking the positive stimulus and reinforcement is longer. (d) It is possible that different parameters characterize the learning rates of different species (19). Perhaps these parameters can be related to the primary modalities of species in learning to adapt to the environment.

Did the ducklings develop a concept of a triangle? The facility with which they chose forms most closely resembling a closed triangle is striking (see p. 165 above). In terms of traditional formulations a positive answer is suggested (2). Although the kind of answer chosen is perhaps simply a matter of definition, there are difficulties in an affirmative one. For example,

respectively. The total time in the experimental situations for both constancies was about four hours. It was not used on any other problem.

⁸ The experiments reported in Ref. 1, 4, 7, 18 can be coordinated to the distinction drawn here.

the ducklings failed to select the triangle when dots only were at the apices (Figure III *j*) and when there was a substantial alteration in the figure-ground relationship (Figure III *b*). If the ducklings had formed a concept the reason for failure with these two figures was not apparent (11, p. 667). On the other hand, if they were responding to cues provided by the figures then we might suppose a breakdown in discrimination when the cues are so reduced that a triangle no longer can be perceived as a physical entity.

G. SUMMARY

Three ducklings learned to select a triangle from trapezoids cut from tinplate, discriminate between vertical and horizontal striations, and demonstrated brightness constancy. One of the two ducklings consistently chose solid cardboard triangles by the age of 12 days and outline triangles by the age of 26 days. The second duckling consistently selected outline triangles by the age of 33 days. In general, the choice of triangle was consistent despite significant changes in contour and figure-ground relationships. The early appearance of these perceptual achievements were interpreted as evidence for their autochthonous determination.

REFERENCES

1. FANTZ, R. I. Form preferences in newly hatched chicks. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 422-430.
2. FIELDS, P. E. Studies in concept formation. *Comp. Psychol. Monog.*, 1932, **9**, (No. 2). Pp. 1-70.
3. GELLERMAN, L. W. Form discrimination in chimpanzees and two-year-old children: I. Form (triangularity). *J. Genet. Psychol.*, 1933, **42**, 3-27.
4. HESS, E. H. Space perception in the chick. *Sci. Amer.*, 1956, **195**, 71-80.
5. KARN, W. H., & MUNN, N. L. Visual pattern discrimination in the dog. *J. Genet. Psychol.*, 1932, **40**, 363-374.
6. LASHLEY, K. S. The mechanism of vision: XV. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.*, 1938, **18**, 123-193.
7. MASON, W. A., BLAZEK, N. C., & HARLOW, H. F. Learning capacities of the infant rhesus monkey. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 449-453.
8. MUNN, N. L. Evolution and Growth of Human Behavior. Boston: Houghton Mifflin, 1955.
9. ———. Handbook of Psychological Research on the Rat. Boston: Houghton Mifflin, 1950.
10. NEET, C. C. Visual pattern discrimination in the Macacus rhesus monkey. *J. Genet. Psychol.*, 1933, **43**, 163-196.
11. OSGOOD, C. E. Method and Theory in Experimental Psychology. New York: Oxford Univ. Press, 1953.
12. PASTORE, N. An examination of one aspect of the thesis that perceiving is learned. *Psychol. Rev.*, 1956, **63**, 309-316.
13. ———. Perceiving as innately determined. *Newsletter* (The Graduate School, Long Island University), 1958, **3**, 6-11.

14. ———. Form perception and size constancy in the duckling. *J. of Psychol.*, 1958, **45**, 259-261.
15. RIESEN, A. H., & MELLINGER, J. C. Interocular transfer of habits in cats after alternating monocular visual experience. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 516-520.
16. SMITH, K. U. Visual discrimination in the cat: II. A further study of the cat for visual figure discrimination. *J. Genet. Psychol.*, 1934, **45**, 336-357.
17. TOWE, A. L. A study of figural equivalence in the pigeon. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 283-287.
18. WINSLOW, C. N. Visual illusions in the chick. *Arch. of Psychol.*, 1933, **153**, 1-83.
19. WODINSKY, J., & BITTERMAN, M. E. Discrimination-reversal in the fish. *Amer. J. Psychol.*, 1957, **70**, 569-576.

Department of Psychology
Queens College
Flushing 67, New York

STUDIES OF PARENT-CHILD RELATIONS: A CRITIQUE*

Department of Psychology, Hofstra College

ROBERT PLUTCHIK AND ESTHER KRONOVET¹

A. PROBLEM

A problem of great theoretical and practical interest concerns the nature of the influence of parental behavior on the formation of the personality of children. Although there has been a tremendous volume of literature dealing with child-raising practices, much of it simply reflects the prevailing customs and mores of a particular society or social class, and is not based on any systematic knowledge of the personalities which result from such practices.

Another source of observation dealing with the relation between parental behavior and childhood personality comes from the many clinical observations of disturbed patients through psychotherapy or interview. These reports have been very valuable in directing attention to particular factors—to certain types of parental behavior—as significant in forming the personality of children, but unfortunately, have not indicated consistent findings. In addition, the statistical studies reported in the literature have been far from unanimous in identifying relevant variables and in describing their influence.

When such a condition exists in science, the sources of the disagreements frequently relate to inadequate methodology. To demonstrate this point and to suggest possible steps for further development will be the purpose of this paper; it will therefore attempt a critical analysis of the methodology of studies of parent-child relationships. These problems will be described under three main headings and will be illustrated by reference to various studies which have been reported in the literature.

B. THE PROBLEM OF THE MEASURING TOOLS USED

Studies of parent-child relations have dealt with one or more of three general questions: (a) personality similarities between parents and children; (b) specific child raising practices in relation to parental attitudes and child behavior; and (c) relations between personality and certain family connected variables such as birth-order, family size, or the presence of broken homes.

* Received in the Editorial Office on May 19, 1958.

¹ This is a revised version of a paper presented at the Fourth Inter-American Congress of Psychology in Puerto Rico, December, 1956.

In each case the personality of the parent and/or the child must be measured and this has been done by various types of tests including multiple-choice (8), sociometric (3), inventory (1) and occasionally projective tests (3). Many of the tests used were developed for the specific study and reliability data is seldom given.

In addition, the dimensions of analysis sometimes seem either irrelevant to the population studied or arbitrary or poorly defined. For example, the *MMPI*, whose terms were developed for use with a psychotic population, was applied to a high school population (2); and terms like "honesty," "loyalty," and "moral courage," have been "measured" by other investigators without adequate definition. In cases where the personality of parents is to be compared with children for similarities, the same inventories are not generally applicable to both. At present there is little evidence on the question of whether a trait found in a child such as "honesty" is the "same" as the trait found in an adult or the extent to which such a relation changes with an increase in the age of the child.

In studies in which ratings or interviews are used, it is especially important to determine the reliability of the judgments. For example, a group of mothers were interviewed by two independent judges on four variables which had already been standardized. The reliability coefficients between the two judges were found to vary from only .46 to .81 on the different scales (4). In contrast is another study in which the mothers of a group of children were visited once by a social worker, after which judgments of unknown reliability were made of such categories as "anxious disposition," or "worried by social conditions" (10).

These examples illustrate the fact that the measuring devices used in studies of parent-child relations have often been inadequate on the grounds of their low or unknown reliability, their arbitrary nature, and their lack of comparability in children and adults. Some recommendations will be made later.

C. THE PROBLEM OF THE BASIC CONCEPTS USED

Much has been written about the effect of "democratic" or "permissive" upbringing on children and yet the published evidence is far from unanimous or conclusive.

This is related to several important factors. The first is that the basic terms used have been differently defined by different investigators. One author, for example, describes "overprotection" simply as babying and too much mothering (6), while another distinguishes "overindulgent overpro-

tection" from "dominating overprotection," in terms of the differing effects on the child (7). Other authors tie the concept of "overprotection" to the notion of "rejection," which is described in such terms as the following: "by rejection is meant a basically hostile attitude overlaid by one which is superficially kind or affectionate. The detection of such an attitude is a matter of inference rather than of direct observation" (5, 15). It is hardly likely that these different ways of measuring or defining what is supposedly the same term will lead to consistent and comparable results.

This can be illustrated by two different studies of "permissiveness"; one defined it by a 20-item attitude scale in which the items were defined *a priori* as restrictive or permissive (1), and the other in which the condition was evaluated after an interview in a person's home (16). The first study concluded that "permissive" and "restrictive" military officers differed in personality; the second, that parents could be permissive at one phase of training and restrictive at another, the term, therefore, not being unequivocally applicable.

A second key problem connected with the basic concepts used is the fact that many of them are value-laden words or ethical terms which tend to prejudice the outcome. Thus, in a study in which "democratic" parents were compared with "authoritarian" parents, the results did not fit the hypothesis that "democratic" parents have better adjusted children; the author therefore discounts his results: "In spite of the (negative) results of this study, the hypothesis of a positive relationship between maternal behavior and the child's early adjustment to school still seems tenable" (4).

Similarly, in another study which expected, but did not find, a linear relation between "security" in college students and breast feeding, the authors conclude: "It was not so much what was done as the spirit in which it was done" (9).

Several other studies have not found a necessary relation between being permissive or liberal in attitude and having well adjusted children (11, 14).

Even more objectionable in scientific research is the use of words like "good" or "bad," "superior" or "inferior," to describe parent-child relations (1, 5). Such ethical terms tend to prejudice the outcome of experiments as well as obscure the fact that the relations between parents and children are extremely complex, never being all-or-nothing, never democratic or authoritarian, good or bad. It is necessary to think in terms of continuous variables rather than polar opposites; it is also necessary to explore the various dimensions (or subscales) which comprise complex, global attitudes such as permissiveness or rejection.

D. THE PROBLEM OF SAMPLING AND GENERALIZATION

Studies of parent-child relations which simply provide some statistics about a special population are of very limited value in understanding possible causal factors. For example, one study reports that more than half of a group of male schizophrenics were single; in more than half their parents had died before the onset of the illness, and so on (13). This type of description, using other measures, has been reported for the mothers of schizophrenics, and for disturbed children in guidance clinics (12, 18, 19). Such information is meaningful only in comparison to other groups which act as controls for the influence of certain variables.

Another problem is related to the fact that many studies deal with special populations, such as upper middle class nursery school children, or college students. The conclusions from such experiments are limited to an unknown degree by the lack of representativeness of the population studied.

One further aspect of this question concerns the sampling of parents. For example, one experiment (17) obtained the attitudes of 1,200 adolescents by a questionnaire and an attitude test was then sent to their parents by mail. Only about 16 per cent of the parents responded, a result which is more or less typical of mail questionnaires, and one which raises the question of the extent to which the parents who responded are typical of the whole group. It is clear that care must be exercised in selecting populations to be studied and in generalizing from them.

E. SUMMARY AND PROPOSALS

It has been suggested in this paper that the many inconsistencies reported in the study of parent-child relations are related to three broad problems: inadequately defined measuring tools; the use of ambiguous or ethically loaded concepts; the use of non-representative groups with resulting over-generalizations.

One important step toward the solution of such problems is to rely more directly on general personality theories to guide research. Such theories must be able to suggest the basic dimensions of family interaction in neutral, descriptive terms which will then provide the basis for the development of tests of high reliability as research tools. These dimensions will be continuous variables, not polar opposites or black and white dichotomies.

The work of Whiting and Child (20), in our opinion, by combining psychoanalytic and reinforcement theory illustrates the approach suggested here. Their analysis enables a comparison of a large number of preliterate

and industrial societies in terms of a few basic dimensions to both test and extend theories.

The development of an integrated body of theory to guide research, as well as a more careful attention to research problems of the kind described here, should have important consequences for the understanding of personality, as well as for the practical guidance of educators and parents as well.

REFERENCES

1. BLOCK, J. Personality characteristics associated with father's attitudes toward child-rearing. *Child Devel.*, 1955, **26**, 41-48.
2. GJERDE, C. M. Parent child resemblances in vocational interests and personality traits. Unpublished Ph.D. thesis, Univ. Minnesota, 1949.
3. HAVIGHURST, R. J., & TABA, H. *Adolescent Character and Personality*. New York: Wiley, 1949.
4. HIGHBERGER, R. The relationship between maternal behavior and the child's early adjustment to nursery school. *Child Devel.*, 1955, **26**, 49-61.
5. HOLMAN, P. Some factors in the aetiology of maladjustment in children. *J. Ment. Sci.*, 1953, **99**, 654-688.
6. KASANIN, J., KNIGHT, E., & SAGE, P. The parent-child relationship in schizophrenia. *J. Nerv. & Ment. Dis.*, 1934, **79**, 249-263.
7. LEVY, D. Maternal overprotection and rejection. *Arch. Neurol. & Psychiat.*, 1931, **25**, 886-889.
8. LYLE, W. H., JR., & LEVITT, E. E. Punitiveness, authoritarianism, and parental discipline of grade school children. *J. Abn. & Soc. Psychol.*, 1955, 42-46.
9. MASLOW, A. H., & SZILAGYI-KESSLER, I. Security and breast feeding. *J. Abn. Psychol.*, 1946, **41**, 83-85.
10. NEUSTATTER, W. L. The effect of poor social conditions in the production of neuroses. *Lancet*, 1938, **134**, 1436-1441.
11. PATTERSON, C. H. The relationship of Bernreuter scores to parent behavior, child behavior, urban-rural residence, and other background factors in 100 normal adult parents. *J. Soc. Psychol.*, 1946, **24**, 3-49.
12. PHILLIPS, E. L. Parent-child similarities in personality disturbances. *J. Clin. Psychol.*, 1951, **7**, 188-190.
13. PLANK, R. The family constellation of a group of schizophrenic patients. *Amer. J. Orthopsychiat.*, 1953, **23**, 817-825.
14. READ, K. Parents expressed attitude and children's behavior. *J. Consult. Psychol.*, 1945, **9**, 96-100.
15. REICHARD, S., & TILLMAN, C. Patterns of parent-child relationships in schizophrenia. *Psychiatry*, 1950, **13**, 247-257.
16. SEWELL, W. H., MUSSEN, P. H., & HARRIS, C. W. Relationships among child training practices. *Amer. Sociol. Rev.*, 1955, **20**, 137-148.
17. STOTT, L. Parental attitudes of farm, town, and city parents in relation to certain personality adjustments in their children. *J. Soc. Psychol.*, 1940, **11**, 325-339.
18. TIETZE, T. A study of mothers of schizophrenic patients. *Psychiatry*, 1949, **12**, 55-65.
19. WAHL, C. W. Some antecedent factors in the family histories of 392 schizophrenics. *Amer. J. Psychiat.*, 1953-54, **110**, 668-676.

20. WHITING, J. W. M., & CHILD, I. L. Child Training and Personality. New Haven: Yale Univ. Press, 1953.

Department of Psychology

Hofstra College

Hempstead, New York

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1959, 95, 177-179.

SEX DIFFERENCES IN THE INDEPENDENCE TRAINING ATTITUDE—*n* ACHIEVEMENT RELATIONSHIP*

Department of Psychology, University of Michigan

JESSE E. GORDON

McClelland notes (1) that whereas high *n* Achievement in boys is associated with early independence training attitudes in their mothers, the reverse may be true for girls. That is, high *n* Achievement in girls may be a consequence of late independence training attitudes in their mothers. The purpose of this paper is to suggest that this position is in error.

McClelland presents some evidence supporting his hypothesis regarding the genesis of *n* Achievement in girls. He recognizes that his *N* of six cases is too small, and his level of significance (.15) too high to provide much support for his position. The criticism presented here is not, however, concerned with the inadequacy of the empirical data.

We proceed from the assumption that independence training attitudes represent antecedent conditions, of which levels of achievement motivation are the consequents. Although it is true that the strength of the relationships between antecedent and consequent conditions often varies in psychology as a function of sex, it is also true that this tends to become more marked in those cases where performance may be influenced by physiological or structural differences between the sexes. Achievement motivation does not appear to describe the kind of behavior which is so influenced. Further, where sex differences affect the relationship strength, they seldom if ever produce a change in the *direction* of the relationship. To suggest that this may occur may be somewhat like suggesting that performance may increase as a function of rewarded trials in boys, but decrease in girls.

It may be true that girls do not respond to independence training attitudes in the same manner as do boys, but if this were so, it would suggest that some other variable which operates in one sex but not the other is at work. That is, if for any given level of an antecedent, *Ss* differ in their responses, then one must look for some other variable as the source of the difference,

* Received in the Editorial Office on April 18, 1958.

rather than assume that the difference represents some true change in the antecedent-consequent relationship.

To illustrate this point further, one may perform a "thought experiment." Suppose that one wished to test the hypothesis that chimpanzees restricted to spoon-feeding by an experimenter would learn to feed themselves later than chimpanzees permitted to practice feeding themselves. One would hardly expect to find a sex difference in the chimpanzees, such that males do learn to feed themselves earlier, but females later, when permitted to practice self-feeding. Although mean age at which self-feeding occurs may differ in males and females, the direction of the relationship between practice and performance remains the same. It is suggested that this "thought experiment" bears some similarity to the independence training-achievement motive relationship studied by McClelland.

If this reasoning is correct, it should be expected that any given level of independence training attitudes should be associated with about the same level of achievement motivation in girls as in boys. If girls have lower *n* Achievement, then one ought to suspect that independence training attitudes favor later independence in girls.

One can identify two sources of the error McClelland has made. First, it is noted that *n* Achievement in males is measured from responses to four *TAT*-like stimuli, all of which represent male figures. It is likely that males project onto the like-sexed figures in the pictures, or identify with them. However, *n* Achievement is measured in females from their responses to the same stimuli. For females, there are no like-sexed figures on whom they may project, or with whom they may identify. It is suggested that the measurement situation is such that, through projection and/or identification, males give personally relevant responses, whereas females may be responding in terms of learned attitudes or expectations regarding the behavior of men. Only indirectly are such responses personally relevant, and the relevance is of a somewhat different kind from that referred to in males.

The second point is with reference to McClelland's finding that *n* Achievement may be aroused in boys by the production of a set to compete with a standard of excellence, whereas this set does not produce the same effect in girls. On the other hand, *n* Achievement in girls may be aroused by a set referring to *S*'s popularity or social acceptance. This suggests that nomologically, the *n* Achievement motive is embedded in substantially different "nets" for boys and girls. We would suggest that where this is so, the concepts cannot be considered to be identical. What is referred to as *n* Achievement in girls cannot be the same as what is referred to as *n* Achievement

ment in boys. It is possible that the measurement procedures used with girls are measuring something more like conformity to social ideals; if this were the case, it would not be difficult to see why scores would vary directly with age of independence training.

In conclusion, what appear to be differences in the relationship between parental independence training attitudes and achievement motivation in boys and girls may result from using the same label, "Achievement, to refer to what may really be two different concepts.

REFERENCES

1. McCLELLAND, ATKINSON, CLARK, AND LOWELL. The Achievement Motive. New York: Appleton-Century-Crofts, 1953. (Author supply complete initials of these authors.)
2. CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. In Feigl and Scriven, *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*. Minneapolis: Univ. Minnesota Press, 1956. (Pp. 174-204.) (Author supply complete initials of these authors.)

Department of Psychology
University of Michigan
Ann Arbor, Michigan

BOOKS

Now that there is a special *APA* journal completely devoted to the publication of book reviews, it is no longer necessary that other journals emphasize such publication. It has always been our conviction that book reviews are a secondary order of publication unless they carry information that is equally important as the book. However, the publication of book titles is a very important service, and we shall continue to render that service.

In any given issue of this journal, we may continue to publish one or more book reviews, but we do not consider such publication a major function of this journal. In line with this policy, we can no longer pay for such manuscripts.

BOOKS RECENTLY PUBLISHED

- ADLER, A. *The Education of the Individual*. New York: Philosophical Library, 1958. Pp. 143.
- BAILEY, E. *Discovering Music with Young Children*. New York: Philosophical, 1958. Pp. 119.
- BALINT, M. *Thrills and Regressions*. New York: International Univ. Press, 1959. Pp. 148.
- BAKER, O. H. *Human Nature Under God: Adventure of Personality*. New York: Association Press, 1958. Pp. 316.
- BARNES, H. E., & TEETERS, N. K. *New Horizons in Criminology*. Englewood Cliffs: Prentice-Hall, 1959. Pp. 654.
- BAY, C. *The Structure of Freedom*. Stanford: Stanford Univ. Press, 1958. Pp. 419.
- BECKER, E. R. *Dictionary of Personnel and Industrial Relations*. New York: Philosophical, 1958. Pp. 366.
- BINDRA, D. *Motivation*. New York: Ronald Press, 1959. Pp. 361.
- BLOCH, H., & NIEDERHOFFER, A. *The Gang: A Study in Adolescent Behavior*. New York: Philosophical, 1958. Pp. 231.
- BONNER, H. *Group Dynamics*. New York: Ronald Press, 1959. Pp. 531.
- BOSSARD, J. H. S., & BOLL, E. S. *Why Marriages Go Wrong*. New York: Ronald Press, 1958. Pp. 224.
- BRAZIER, M. A. B., Ed. *The Central Nervous System and Behavior*. New York: Josiah Macy, Jr. Foundation, 1959. Pp. 450.
- BROADBENT, D. E. *Perception and Communication*. New York: Pergamon Press, 1958. Pp. 338.
- BROWER, D., & ABT, L. E., Eds. *Progress in Clinical Psychology*. New York: Grune & Stratton, 1958. Pp. 249.
- BROWN, N. C. *Life Against Death*. Middletown: Wesleyan Univ. Press, 1959. Pp. 366.
- BROWN, R. *Words and Things*. Glencoe: Free Press, 1958. Pp. 398.
- BROWNE, C. C., & COHN, T. S., Eds. *The Study of Leadership*. Danville: Interstate Printers & Publishers, 1958. Pp. 487.
- CANTRIL, H. *The Politics of Despair*. New York: Basic Books, 1958. Pp. 269.
- CARTER, R. *The Doctor Business*. Garden City: Doubleday, 1958. Pp. 283.
- CHESSER, E. *Live and Let Live*. New York: Philosophical, 1958. Pp. 125.
- CHOLST, S. *The Only Baby*. New York: Whittier, 1958. Pp. 71.
- COLBY, K. M. *A Skeptical Psychoanalyst*. New York: Ronald Press, 1958. Pp. 145.
- COLE, L. E., & BRUCE, W. F. *Educational Psychology*. Yonkers-on-Hudson: World Book, 1958. Pp. 701.
- COONEY, T. *Ultimate Desires*. New York: Philosophical Library, 1958. Pp. 99.
- COX, D. R. *Planning of Experiments*. New York: Wiley, 1958. Pp. 308.

- DAVIES, S. P. *The Mentally Retarded in Society*. New York: Columbia Univ. Press, 1959. Pp. 248.
- DE MANN, R. F., & HAVIGHURST, R. J. *Educating Gifted Children*. Chicago: Univ. Chicago Press, 1957. Pp. 276.
- DOWNIE, N. M. *Fundamentals of Measurement*. New York: Oxford Univ. Press, 1958. Pp. 413.
- EKSTEIN, R., & WALLERSTEIN, R. S. *The Teaching and Learning of Psychotherapy*. New York: Basic Books, 1958. Pp. 334.
- EMERY, F. E., & OESER, O. A. *Information, Decision, and Action*. Melbourne: Melbourne Univ. Press, 1958. Pp. 132.
- ERIKSON, E. H. *Young Man Luther*. New York: Norton, 1958. Pp. 288.
- FALK, D. V. Eugene O'Neill and the Tragic Tension. New Brunswick: Rutgers Univ. Press, 1958. Pp. 211.
- FEIN, R. *Economics of Mental Illness*. New York: Basic Books, 1958. Pp. 164.
- FLEMING, C. M. *Teaching: A Psychological Analysis*. New York: Wiley, 1958. Pp. 291.
- GALLAGHER, J. R., & HARRIS, H. I. *Emotional Problems of Adolescents*. New York: Oxford Univ. Press, 1958. Pp. 174.
- GOLDBRUMNER, J. *Cure of Mind and Cure of Soul*. New York: Pantheon, 1958. Pp. 127.
- GORDON, H. L., *Ed.* *The New Chemotherapy in Mental Illness*. New York: Philosophical, 1958. Pp. 762.
- GROSS, E. *Work and Society*. New York: Crowell, 1958. Pp. 652.
- GUILFORD, J. P. *Personality*. New York: McGraw-Hill, 1959. Pp. 562.
- HADFIELD, J. A. *Psychology and Morals*. New York: McBride, 1958. Pp. 219.
- HAUSSERMANN, E. *Developmental Potential of Preschool Children*. New York: Grune & Stratton, 1958. Pp. 285.
- HAGGARD, E. A. *Intraclass Correlation and the Analysis of Variance*. New York: Henry Holt, 1958. Pp. 171.
- HAKKINEN, S. *Traffic Accidents and Driver Characteristics*. Helsinki: Akateeminen Kirjakauppa, 1958. Pp. 198.
- HENRY, H. *Motivation Research*. New York: Ungar, 1958. Pp. 240.
- HIMMELWEIT, H. T., OPPENHEIM, A. N., & VINCE, P. *Television and the Child*. New York: Oxford Univ. Press, 1958. Pp. 522.
- HUGHES, E. C. *Men and Their Work*. Glencoe: Free Press, 1958. Pp. 184.
- JAHODA, M. *Current Concepts of Positive Mental Health*. New York: Basic Books, 1958. Pp. 136.
- JANIS, I. L. *Psychological Stress*. New York: Wiley, 1958. Pp. 439.
- KITTELLE, P. W. *Voyage to the Unknown*. New York: Vantage, 1958. Pp. 132.
- KNELLER, G. F. *Existentialism and Education*. New York: Philosophical, 1958. Pp. 170.
- KOCH, S. *Psychology: A Study of a Science*. New York: McGraw-Hill, 1959. Pp. 710.
- KUBIE, L. S. *Neurotic Distortion of the Creative Process*. Lawrence: Univ. Kansas Press, 1958. Pp. 151.
- KUPER, L., WATTS, H., & DAVIES, R. *Durban: A Study in Racial Ecology*. New York: Columbia Univ. Press, 1958. Pp. 254.
- LARRABEE, E., & MEYERSON, R., *Eds.* *Mass Leisure*. Glencoe: Free Press, 1958. Pp. 460.
- LEAVITT, J. E., *Ed.* *Nursery-Kindergarten Education*. New York: McGraw-Hill, 1958. Pp. 365.
- LIEBMAN, S., *Ed.* *Emotional Problems of Childhood*. Philadelphia: Lippincott, 1958. Pp. 176.
- LINDZEY, C., *Ed.* *Assessment of Human Motives*. New York: Rinehart, 1958. Pp. 273.
- MAGOUN, H. W. *The Waking Brain*. Springfield: Thomas, 1958. Pp. 138.
- MAIER, N. R. F. *The Appraisal Interview*. New York: Wiley, 1958. Pp. 246.
- MARCH, J. G., & SIMON, H. A. *Organizations*. New York: Wiley, 1958. Pp. 262.

- MARIANO, J. H. *The Use of Psychotherapy in Divorce and Separation Cases*. New York: American Press, 1958. Pp. 179.
- MAXWELL, A. E. *Experimental Design in Psychology and the Medical Sciences*. New York: Wiley, 1958. Pp. 147.
- McDANIEL, E. F. *Discovering the Real Self*. New York: Philosophical, 1958. Pp. 117.
- McEWEN, P. *Figural After-Effects*. New York: Cambridge Univ. Press, 1958. Pp. 106.
- MERTON, R. K., BROOM, L., & COTTRELL, L. S., JR. *Sociology Today*. New York: Basic Books, 1959. Pp. 623.
- MUDRY, J. *Philosophy of Atomic Physics*. New York: Philosophical, 1958. Pp. 136.
- MURPHY, G. *Human Potentialities*. New York: Basic Books, 1958. Pp. 340.
- NAFZIGER, R. O., & WHITE, D. M., *Eds.* *Introduction to Mass Communications Research*. Baton Rouge: Louisiana State Univ. Press, 1958. Pp. 244.
- NYE, F. I. *Family Relationships and Delinquent Behavior*. New York: Wiley, 1958. Pp. 168.
- PIERCE, J. R., & DAVID, E. E. *Man's World of Sound*. Garden City: Doubleday, 1958. Pp. 287.
- PEARSON, G. H. J. *Adolescence and the Conflict of Generations*. New York: Norton, 1958. Pp. 186.
- PHELPS, W. M., HOPKINS, T. W., & COUSINS, R. *The Cerebral-Palsied Child*. New York: Simon & Schuster, 1958. Pp. 237.
- PREU, J., *Ed.* *The Negro in American Society*. Tallahassee: Florida State Univ., 1958. Pp. 89.
- ROE, A., & SIMPSON, G. G., *Eds.* *Behavior and Evolution*. New Haven: Yale Univ. Press, 1958. Pp. 557.
- ROSEN, G. *Industrial Change in India*. Glencoe: Free Press, 1958. Pp. 243.
- ROUCEK, J. S., *Ed.* *Contemporary Sociology*. New York: Philosophical, 1958. Pp. 1209.
- ROTHNEY, J. W. M., DANIELSON, P. J., & HEIMANN, R. A. New York: Harper, 1959. Pp. 378.
- SAKEL, M. *Epilepsy*. New York: Philosophical, 1958. Pp. 204.
- SAKEL, M. *Schizophrenia*. New York: Philosophical, 1958. Pp. 335.
- SALIT, C. R. *Man in Search of Immortality*. New York: Philosophical, 1958. Pp. 185.
- SAUL, L. J. *Technic and Practice of Psychoanalysis*. Philadelphia: Lippincott, 1958. Pp. 244.
- SAYLES, L. R. *Behavior of Industrial Work Groups*. New York: Wiley, 1958. Pp. 182.
- SENDERS, V. L. *Measurement and Statistics*. New York: Oxford Univ. Press, 1958. Pp. 594.
- SEWARD, C., *Ed.* *Clinical Studies in Culture Conflict*. New York: Ronald Press, 1958. Pp. 598.
- SHELDON, H. D. *The Older Population of the United States*. New York: Wiley, 1958. Pp. 223.
- SLAVSON, S. R. *Child-Centered Group Guidance of Parents*. New York: International Univ. Press, 1958. Pp. 333.
- SMITH, P., *Ed.* *Creativity*. New York: Hastings House, 1959. Pp. 210.
- STACEY, C. L., & DE MARTINO, M. F., *Eds.* *Understanding Human Motivation*. Cleveland: Howard Allen, 1958. Pp. 507.
- STOGDILL, R. M. *Individual Behavior and Group Achievement*. New York: Oxford Univ. Press, 1959. Pp. 352.
- STRENG, A., FITCH, W. J., HEDGECOCK, L. D., PHILLIPS, J. W., & CARRELL, J. A. *Hearing Therapy for Children*. New York: Grune & Stratton, 1958. Pp. 353.
- TANNER, J. N., & INHELDER, B., *Eds.* *Discussions on Child Development*. New York: International Univ. Press, 1958. Pp. 223.
- TUMIN, M. M. *Desegregation*. Princeton: Princeton Univ. Press, 1958. Pp. 270.
- ULLMANN, S. *The Principles of Semantics*. New York: Philosophical, 1957. Pp. 346.
- WATSON, R. I. *Psychology of the Child*. New York: Wiley, 1959. Pp. 662.

\$14.00 per year
\$10.00 per volume
Single Numbers \$5.00

QUARTERLY
Two volumes per year

December, 1959
Volume 95, Second Half

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

THE JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior,
and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

DECEMBER, 1959

(Manuscripts are printed in the order of final acceptance)

Some effects of a low protein diet on a first filial generation of white rats	187
By J. J. COWLEY AND R. D. GRIESEL	
A method for appraising children in in-patient psychiatric treatment: Ranking for normality	203
By WILLIAM GOLDFARB	
Partial reinforcement following a long continuous reinforcement acquisition series	207
By DONALD J. LEWIS AND DONALD A. HARDER	
A scale to identify impulsive behavior in children	211
By B. SUTTON-SMITH AND B. G. ROSENBERG	

(OVER)

Copyright, 1959, by The Journal Press
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter May 11, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879

The effect of stimulant and depressant drugs on a measure of emotional reactivity in the rat	217
By P. L. BROADHURST, S. N. SINHA, AND S. D. SINGH	
The expression of guilt in fantasy and reality	227
By KENNETH R. WURTZ	
An assessment of the fakability of scores on the group Personality Projective Test	239
By RUSSELL N. CASSEL AND ROBERT P. BRAUCHLE	
Manipulation of visually homogeneous stimulus objects	245
By RICHARD M. CARR AND W. LYNN BROWN	
Observational learning of an avoidance response	251
By W. J. PRESLEY AND A. J. RIOPELLE	
Effect of exogenous adrenalin on the startle response in the rat	255
By K. E. MOYER AND B. N. BUNNELL	
The effects of early enforced weaning on sucking behavior of puppies	261
By J. P. SCOTT, SHERMAN ROSS, AND ALAN E. FISHER	
Effects of differential instructions and anxiety level on discrimination learning	283
By ROBERT E. MINNICH AND PERRY LONDON	
Exploratory studies of verbal context by means of clustering in free recall	293
By RICHARD C. GONZALEZ AND CHARLES N. COFER	
The effects of food deprivation and restricted activity upon exploratory behavior of the rat	321
By R. M. CARR, J. E. OVERALL, R. K. WHITE, AND W. L. BROWN	
BOOKS	
Sargant, W. <i>Battle for the Mind: A Physiology of Conversion and Brain-Washing.</i> (Reviewed by Gladys C. Schwesinger)	329
Anastasi, A. <i>Differential Psychology</i> , 3rd Ed. (Reviewed by Gladys C. Schwesinger)	333
BOOKS RECENTLY PUBLISHED	337

SOME EFFECTS OF A LOW PROTEIN DIET ON A FIRST FILIAL GENERATION OF WHITE RATS*

National Institute for Personnel Research, Johannesburg

J. J. COWLEY AND R. D. GRIESEL¹

A. INTRODUCTION AND PROBLEM

The investigation forms part of a study by this Institute² into the effects of malnutrition on the physical and mental development of the African people. A high incidence of malnutrition and of malnutritional disorders is a feature of many of the African territories. Kwashiorkor, one of the most widespread of the nutritional disorders in Africa, is generally accepted as being due to a deficiency of protein. It is more common in childhood and is particularly associated with the period of weaning.

While no clear indication of the extent of the nutritional disorders in Africa can be given, the Kwashiorkor syndrome has been described in West, East, Central, and South Africa. The reported incidence of the syndrome varies from one area to another and probably reflects on the degree of development of the area and also on differences in diagnosis. In one region of Dakar, Raoult (17) describes the syndrome as being found in 10 to 20 per cent of children of weaning age. Squires (23) gives the mean incidence of deficiency diseases as 2 per cent for the Bechuanaland Protectorate but emphasizes that the figure represents only the most serious cases. Elsewhere, Squires (22) reports that field surveys of school children in the larger settlements of Bechuanaland showed an incidence of between 20 and 30 per cent with nutritional deficiencies.

The extent to which nutritional deficiencies influence the effectiveness and achievement of a community is unknown, but they are regarded by many as being an important retarding influence. In an inaugural address to the Second Inter-African Conference on Nutrition Sir Percy Wyn-Harris (25), Governor of the Gambia, stressed the view that Kwashiorkor could well be

* Recommended for publication by S. Biesheuvel, received in the Editorial Office on August 1, 1959, and given immediate publication because of its implications for cross-cultural research in South Africa.

¹ The authors wish to express their appreciation to the National Nutrition Research Institute for their advice and assistance with the diets, and to the University of Natal for the facilities provided by the Department of Psychology.

² Developmental Research Unit of the National Institute for Personnel Research, C.S.I.R., at the University of Natal.

responsible for reducing the working capacity of entire African communities. A similar view has been expressed by Brock and Autret (6) who attribute the backwardness of the African people partially to a dietary deficiency of protein.

The present study represents an attempt to establish the effect of a low protein diet on the intelligence of albino rats reared from mothers that have themselves been fed a low protein diet.

A number of studies have drawn attention to the effects of low protein diets on maze learning in rats. Riess and Block (18) reported impairment in maze learning in rats on low cystine and lysine diets. The rats were placed on the low protein diets at the age of 31 days.

Bernhardt (3) used an escape from water maze and found consistent differences between a low protein group of rats and a control group. The low protein rats were poorer in maze learning than the control though the observed differences were not at an acceptable level of statistical significance.

Bevan and Freeman (5) also used a diet low in lysine but found no significant differences in maze learning between the group of rats maintained on the low protein diet from the age of 30 days and the control group fed on a standard laboratory diet. The control group was restricted in amount of food to the mean weight consumed by the experimental group. Bevan and Freeman did find, however, that the relearning of the maze was reliably more efficient in the case of the control animals.

Rather different findings are reported by Pilgrim *et al.* (15). Using an escape from water maze, Pilgrim reported no significant differences between a group of rats maintained on a low protein diet, a group on a high protein diet, and a group on a basal diet. The rats were fed on the low protein diet from about the 34th day after birth and were kept on the diet for 40 to 50 days.

In an early study, Anderson and Smith (1) compared the maze learning performance of rats fed a low lysine diet with a group of rats maintained on a restricted diet of "insufficient calorie intake" and with a group on a control diet. In the main experiment, the experimental diets were introduced when the rats were 44 days of age. The rats stunted in growth by the low protein diet and by the restricted diet are reported as being superior to normal rats in relearning a maze. The error scores of the stunted rats tended to be higher than those of the control group, but there was little consistency in the pattern of error scores and differences tended to be eliminated after a period of realimentation.

There is little consistency in the studies on the effect of low protein diets on maze learning. The low protein diets have been fed to young rats at different ages and for varying lengths of time. The behaviour of the rats which has been recorded is limited, for the main part, to observations on their performance on a number of different types of mazes and to an analysis of their activity records (4, 21).

The present report describes the effects of a low protein diet on intelligence, exploratory behaviour, and growth in a first filial generation of white rats. A further study aimed at investigating the effects of the low protein diet on further filial generations is in progress.

B. METHOD

1. *The Parent Generation*

An experimental group of 10 female albino rats was fed from 21 days of age on a diet low in protein. A comparable control group of 10 female rats was retained on the laboratory diet. At 90 days, the two groups were mated with male rats from the same Wistar strain. Both groups of male rats had been fed on the laboratory diet. The experimental female rats and the control female rats are referred to as the parent generation. The observations in this paper will primarily refer to the first filial generation offspring of these two parent groups. Some observations on the parent generation are, however, included.

2. *The First Filial Generation*

The experimental group of the first filial generation was bred and reared, as described, from mothers who had themselves been fed on the low protein diet. The first filial generation of control rats was bred and reared in the same way as the experimental rats, except that they were retained throughout on the standard laboratory diet.

The experimental and control animals of the first filial generation were kept under the same laboratory conditions and received the same handling. In both groups, the pups remained with the mother for 35 days after birth. After separation from the mothers, the rats were kept in pairs in cages measuring 30 x 21 x 21 centimeters.

At no time did the experimental animals of the first filial generation have access to any food other than that of the low protein diet. Rats of the parent generation and of the first filial generation had unrestricted access to their diet and to water throughout the experiment.

3. *Diets Used in the Investigation*

The composition of the low protein diet was based on the work of Miller and Platt (14). Miller and Platt compounded a basal diet in which the foodstuffs, though of European origin, corresponded in amount to those of a rural area in Gambia. Varying amounts of protein in the form of fish meal were added to the basal diet. The reproductive performance of albino rats maintained on the different diets was observed. They found that with a diet of 5 per cent fish meal they were able to wean 50 per cent of the rat litters born, while with a diet of 2 per cent fish meal only 33 per cent of the litters born could be weaned.

On account of differences in the composition of local foodstuffs, it was necessary in this study to make changes in the amounts of some of the foodstuffs. Thus, $7\frac{1}{2}$ per cent fish meal was used in the experimental diet so that the total protein content was 12.97 per cent. Miller and Platt, on the other hand, used 2 per cent fish meal to supplement their basal diet and report a total protein content of 13.8 per cent, as compared to 13.5 per cent for a Gambian diet. The composition of the experimental and control diets is shown in Table 1. It is appreciated that, though the protein content of the diet is known, not all the protein is necessarily available to the rat. The amino acids present in the one diet may have different "biological value" from the amino acids present in the other (6, 7).

4. *Test Material and Procedure*

The apparatus used for assessing the intelligence of the rats was the enclosed field test of Hebb and Williams (11). A description of the test is

TABLE 1
COMPOSITION OF THE LOW PROTEIN AND CONTROL DIETS

Control	%	Low protein	%
Yellow maize meal	56	Ground whole wheat	75
Milk powder	16	Ground cooked haricot beans	5
Ground nuts	6	Fish meal	7
Brewers yeast	5	Dextrine	5
Fish meal	8	Ground nut oil	2
Carcase meal	4	Mineral mixture*	1
Lucerne meal	3	Vitamin mixture*	2
Bone meal	1	Vitamin A and D	.007
Calcium	.5	Lucerne meal	3
Salt	.5		
Vitamin A and D ₃	.1		
Percentage protein	20.1	Percentage protein	12.97

* Mineral and vitamin mixtures supplied by the South African Council for Scientific and Industrial Research, National Nutrition Research Institute (7).

given by the originators with supporting evidence of its validity and reliability. Minor changes in the design of the apparatus and modifications in testing procedure are described below.

The test consists of an enclosure 30 x 36 inches in size with walls of a uniform height of four inches. The enclosure is covered by a detachable wire screen. The starting box and the goal box are in fixed positions in opposite corners of the enclosure.

In the present study the hardboard floor of the enclosure was marked out in 30 squares of equal area, and the barriers for the various problems were erected on the lines demarcating the squares.

The barriers, which are of various lengths, can be erected to form different problems, or, what Hebb and Williams term *Umweg* situations. The problems used in the testing were the same as those described by Hebb and Williams. Each rat was tested on three problems on each of four consecutive days.

There was a preliminary training phase during which no barriers were erected in the enclosure and the food in the goal box was the same as that on which the rat exploring the enclosure had been reared. The rat's route through the enclosure was plotted and the number of squares it entered during a three-minute period was counted. The procedure was repeated each day for 15 days with a break of two days after each period of five days. By the end of the first phase nearly all the rats would run directly to the goal box on entering the enclosure.

During the second phase the barriers forming the *Umweg* situations were erected. The food in the goal box was a mixture of equal portions of the low protein and the control diets. The rats had been deprived of food for 24 hours prior to being tested. The path of the rat was again traced, and the time taken to reach the goal box was recorded. A maximum time of one minute was allowed for each of the 12 problems, and each rat was given only one run on each of the problems. Performance on each of the problems was judged by two independent assessors using a three-point scale. The assessors judged the directness of the run to the goal box, assigning a high score to those rats which used the shortest and most direct path to the goal box.

In testing the rats on the Hebb-Williams test the same number of rats were randomly selected from the control as were available in the experimental group. In the experimental group there were 16 female and 12 male rats and in the control there were 16 male and 12 female rats. The sex-ratio differences are not statistically significant.

5. Test Feeding

The amount of food removed each day from the food containers of eight male experimental and eight male control rats was recorded over a period of 12 days when the rats were 215 days old. The weight of food removed each day from the non-spill containers was regarded as an index of the amount of food eaten. The mean weight of food ingested by the experimental rats was 16.5 grams per day, as compared with 16.8 grams per day for the control rats. The difference in the mean weight of food ingested by the experimental and the control rats is clearly not significant.

C. RESULTS

1. Growth and Reproduction

Mean differences at various ages in the weight of the experimental and the control rats in the parent and first filial generations are shown in Figure 1.

In the parent generation, the mean weight of the control group was significantly greater than that of the experimental group at 200 days ($t = 3.709$, $df = 18$, $p = .01$).

In the first filial generation, mean differences in the weight of the two groups were not appreciable at 28 days after birth, but the difference between

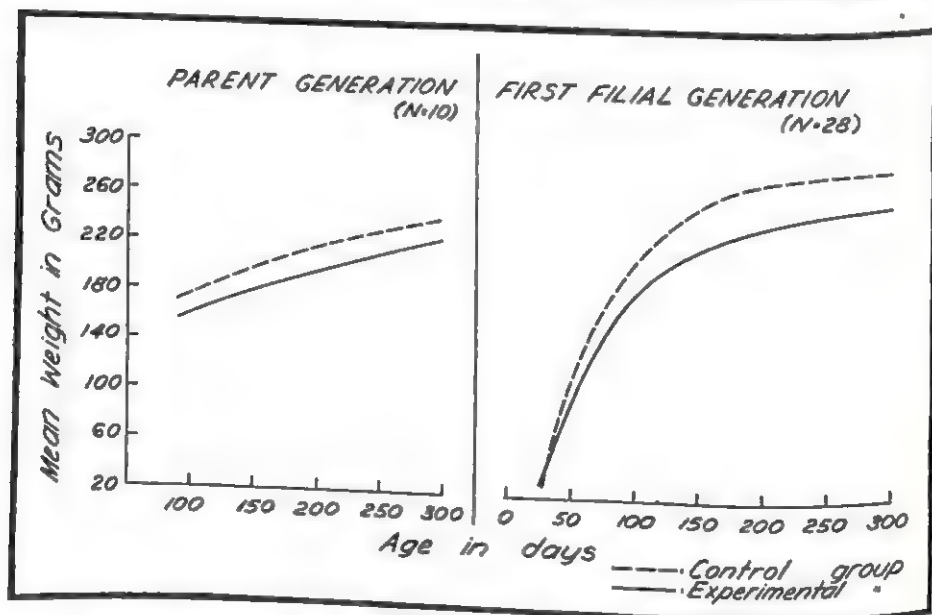


FIGURE 1
WEIGHT CURVES OF THE PARENT AND THE FIRST FILIAL GENERATIONS

the two groups had increased by 4.8 grams at 35 days when the pups were moved from their mothers. At 90 days, the mean difference in the weight of the two groups was 23.96 grams and the difference had shown a further increase to 36.6 grams when the rats were 175 days old. At both 90 and 175 days the difference in the mean weight of the two groups is significant at the five per cent confidence level.

In the experimental group, 47 pups were born as compared with 41 in the control group. The mean number of pups per litter was higher, however, in the control group (mean = 6.8) than in the experimental group (mean = 5.2). In the control group four of the 10 females failed to produce litters, while in the experimental group, three of the 10 females failed to do so. The differences are not statistically significant.

At the time that the first filial generation pups were separated from their mothers, only 63.8 per cent of the experimental pups were living as compared with 95.1 per cent of the pups in the control group. The chi-squared test shows the difference in mortality rate to be highly significant ($\chi^2 = 10.88$, $df = 1$, $p = .001$).

Eighty-eight per cent of the deaths in the experimental pups occurred between the 12th and 24th day after birth. Seventy per cent of the deaths in this group occurred on either the 18th, 19th, or 20th day after birth.

Figures on the number of pups born and on the mortality rate must be treated as approximations. It is likely that the number of pups born was higher than that recorded, but that a number were eaten before their births could be recorded.

From 35 days of age until maturity (70 days), a further two pups died in the experimental group. During this period there were no deaths in the control group. Between maturity and 520 days, nine of the experimental rats died. All nine deaths in the experimental group occurred between the ages of 289 and 520 days. In the same period there were no deaths in the control group ($\chi^2 = 5.15$, $df = 1$, $p = .02$).

2. *Exploratory Behaviour*

As described, the base of the enclosed field test was divided into a number of equal areas by a series of lines painted on the inner surface. The number of areas entered by the rat during the preliminary training phase were counted and regarded as a measure of the rat's exploratory behaviour (active exploration).

In the first filial generation, the mean differences in exploratory behaviour

between the low protein group and the control group is not significant (Table 2). A comparison of the mean scores of the low protein group and the control

TABLE 2
MEAN DIFFERENCES IN EXPLORATORY BEHAVIOUR BETWEEN THE LOW PROTEIN AND CONTROL RATS (F1)

Group	Mean low protein	Mean control	Standard deviation low protein	Standard deviation control	<i>t</i>	<i>df</i>	<i>p</i>
First filial (15 trials)	854.75	961.7	264.0	254.12	1.54	54	Not sig.
First filial (first 8 trials)	411.11	474.21	151.34	146.33	1.59	54	Not sig.
First filial (second 7 trials)	443.64	487.5	129.11	124.90	1.29	54	Not sig.
Parent generation (15 trials)	748.2	699.1	466.8	287.6	.27	16	Not sig.

group over the early (first eight) and last (last seven) training periods also failed to show any significant difference in exploratory behaviour. The mean score of the control group is, however, greater than that of the low protein group during the early and late training periods, as it is also over the combined period of training.

The parent generation was not tested on the enclosed field test except to obtain an index of their exploratory behaviour. The mean difference in exploratory behaviour of the low protein group was greater than that of the control group but the difference between the groups is not significantly different (Table 2).

3. Intelligence

The first filial generation of the low protein rats scored significantly lower on the problems of the Hebb-Williams test than did the control rats. The low protein rats were also significantly slower in reaching the goal box than the control group. The results are shown in Table 3.

Exploratory behaviour did not correlate significantly with performance on the problems of the Hebb-Williams test. The relationship between exploratory behaviour and performance on the Hebb-Williams test is positive in the low protein group ($r = .26$) and negative in the control group ($r = -.21$).

TABLE 3
MEAN DIFFERENCES IN INTELLIGENCE BETWEEN THE LOW PROTEIN AND CONTROL RATS (F1)

Test scoring	Mean low protein	Mean control	Standard deviation low protein	Standard deviation control	<i>t</i>	<i>p</i>
Rater "A"	20.10	23.35	3.59	2.97	3.69	.1%
Rater "B"	18.10	21.0	2.81	2.72	3.91	.1%
Rater "A" + "B"	38.21	44.36	6.20	5.40	3.95	.1%
Time score (seconds)	446.86	365.35	110.98	95.94	2.93	1%

D. DISCUSSION

The results have shown a number of marked differences between the first filial generation fed the low protein diet and the control rats fed the standard laboratory diet.

The parent generation and the first filial generation are lighter in weight than the comparable control groups. Mean weight differences in the first filial generation are observable at weaning, but only reach a significant level at 90 days of age. Brock and Autret (6) regard a retardation of weight as characteristic of Kwashiorkor as well as of other malnutritional conditions. These authors compare accepted weight curves of European infants and children with weight curves of African infants and children who attended clinics in the Kampala area of Uganda. These curves show a striking similarity to the weight curves of Figure 1.

Miller and Platt (14) have reported a marked decrease in the mortality rate of pups reared from mothers on a poor protein diet, when the pups are left to run with the mothers for the fourth and fifth weeks of life. They consider confirmation of this finding will provide strong argument in favour of prolonged breast feeding in regions where diets are marginal or sub-optimal in nutritive value. In the present study, the mortality rate of the low protein pups was high, and showed an unaccountable increase in the third week of life though the pups were not removed from the mothers until they were 35 days of age. It seems likely, in view of Miller and Platt's observations, that the mortality rate of the low protein rats would have been higher had they been removed from their mothers at an earlier age.

Up to the age of 18 months, the low protein adult rats of the first filial generation showed a higher mortality rate than the control rats among whom there were no deaths. At autopsy, lesions of the lungs were found in eight

of the nine low protein rats, though the extent of the lesions varied considerably. In the case of the ninth rat, the lungs had been eaten by a cage mate before the rat was found. Further, five of the nine rats showed signs of either intestinal or gastric lesions. In three of the five rats the intestine was deflated, in a fourth the stomach was blocked by a hair ball, and in the fifth rat the upper intestine was constricted. Histological and bacteriological tests were not conducted and the cause of death cannot be specified. Machella and Griffith (13), quoting a review by Shay *et al.* (19), report that gastric lesions are a common accompaniment of a number of dietary deficiencies including protein and amino acid deficiencies.

In the absence of any deaths during this period in the control group no comparisons were made and the significance of the pulmonary and gastrointestinal lesions is not known. Slonaker (20) investigated the effect of different percentages of protein in the diet of successive generations of rats and recorded that at autopsy lesions were present in the lungs and alimentary tracts of a high percentage of the rats. He considers that the lesions might well be attributed to the dry powdery food as they were found to occur independently of the protein composition of the diet. While the powdery nature of the food may account for the lesions, as suggested by Slonaker, the problem of the significantly higher mortality rate in the low protein group remains, as both low protein and control rats were fed powdered diets.

The low protein rats and the control rats of both the parent and the first filial generation showed no marked difference in their exploratory behaviour. Riess and Block (18) have similarly reported that parent generation rats fed diets low in cystine and lysine show no statistically reliable differences in activity when compared with rats fed a standard laboratory diet. In contrast to these results Bevan *et al.* (4) report that a small group of parent generation rats retained on a low lysine diet ($N = 5$) showed a higher activity level than a control group ($N = 5$) fed the standard laboratory diet.

In repeating the present study on another parent generation of rats it was again found that there was no marked difference in exploratory behaviour between a low protein and a control group of rats. There was also no significant difference in activity as measured on an activity wheel. The results are regarded as confirming the finding of Riess and Block as regards the parent generation and indicate that exploratory behaviour in the first filial generation, also, is not affected by the low protein diet.

Differences in intelligence between the low protein and the control group were marked in the first filial generation. The low protein rats were slower

in reaching the goal box and their performance on the problems of the Hebb-Williams test was inferior to that of the control group. A subsequent study to which reference has already been made, has indicated that in a parent generation the Hebb-Williams test does not discriminate between rats fed the low protein diet and rats fed the laboratory diet. It is possible that the pups have been handicapped by the low protein diet of their mothers either during the period of gestation, or during the period of breast feeding, or that the two periods together have had a cumulative effect. The defects of the mothers' diet appear likely to have aggravated any direct effects that the low protein diet had on the pups.

Occasional signs of malnutrition have been reported in African infants at six and even four months of age (6), while Gillman and Gillman (9) report that African infants are sometimes born with gross signs of malnutrition.

African children are often breast-fed until the age of 18 months or even longer, and the question arises as to whether the quantity, or the quality, of the breast milk, or both, are likely to be affected by the diet of the mother. Espe and Smith (8) report that, in cows, low protein feeding tends to depress the protein and fat content of the milk, but that its most serious effect is in lowered milk production. The work of Kon and Mawson (12) shows that in the case of human milk, there is adequate evidence that the vitamin content of the milk can be directly affected by supplementary feeding. On the other hand, there appears to be considerable divergence of opinion as to the effect of a low protein diet on the quality of the mothers' milk, though many workers appear to accept that quantitative changes may occur. Brock and Autret (6) report that in a number of areas of Africa there is an impression that both the quantity and quality of the milk decline from the age of six months. Auffret and Tanguy (2) report that in Dakar the methionine content of the mothers' milk was lower than that reported for European mothers. Trowell *et al.* (24) consider that if the work of Auffret and Tanguy is substantiated, then a good case can be made out for improving the nutrition of a community by supplementing the mothers' diet during pregnancy.

Brock and Autret (6) have pointed out that the introduction of supplementary feeding at an early age in many African communities does not alleviate the protein deficiency as the supplementary feeds are the ordinary foods of the family and these also are low in protein. The authors further point out that in children the period of 1-4 years is a period of high protein requirement and that this period coincides with a high incidence of Kwashiorkor.

In reviewing the literature on the effects of low protein diets on the be-

haviour of the albino rat, it was noted that there was little consistency in the studies and that the studies were restricted to maze learning in parent generation rats. In these studies the rats were placed on the low protein diets after weaning, and presumably after the period when a low protein diet would have had the greatest effect on the development and behaviour of the rat.

The results obtained in this study have been attributed to the effects of the low protein content of the experimental diet. The possibility that some other difference between the two diets has contributed to, or even been responsible for, the observed differences cannot be entirely excluded. Certain consistencies between the present study and other studies on the effect of low protein diets tend, however, to corroborate the view that it is the low protein content which is responsible for the differences observed. Thus, retardation of growth as represented by weight curves, has been shown by Anderson and Smith (1) and Slonaker (20) to accompany the feeding of a low protein diet to rats. Pollack and Halpern (16) report that loss of weight is a consistent finding in protein deprivation and that the loss of weight parallels the extent of the deficiency. Further support for attributing the results to the low protein content of the diet is to be found in the work of Miller and Platt (14) and Slonaker (20). They have found, as has the present study, that there is a high mortality rate in the period from birth until weaning when rats are retained on a low protein diet.

The choice of a food incentive may be questioned, for the significance of the incentive may vary according to the particular group tested. Griffiths and Senter (10) reported that rats on a protein free diet were superior in maze running to rats on a standard diet when the balanced diet was the incentive. In a similar way, rats on a protein free diet with a balanced diet as incentive were reported as being superior in maze learning to rats on a protein free diet and with a protein free diet as incentive.

Any type of incentive, however, may have a different significance for the control group as compared with the low protein group and this could be the case when the incentive is to escape from water or from an electric shock, as well as when the incentive is to obtain food.

Following Griffiths and Senter, it is possible to assume that an increase in the protein content of the food in the goal box would provide an added incentive for the low protein group. Thus, when the rats were tested on the second phase of the Hebb-Williams test and given a mixture of equal portions of the control and low protein diets, the incentive for the low protein rats would be greater than for the control group. It was found, however, that

the low protein rats were less successful than the control group. This suggests that the impairment consequent upon the low protein diet could not be compensated for by the greater significance of the incentive provided.

E. SUMMARY AND CONCLUSIONS

Malnutrition and malnutritional disorders are widespread throughout the African continent. One of the most widespread of the malnutritional disorders is the Kwashiorkor syndrome. It has been postulated that this syndrome is partly responsible for the backwardness of many of the African people.

In attempting to evaluate this hypothesis it was decided to study the effect of a low protein diet on successive generations of white rats. The present study reports on a number of observations made on a first filial generation.

The low protein diet selected was based on an analysis by Miller and Platt (14) of an African diet of the Gambia. The diet used for the control groups was a standard laboratory diet of known protein composition.

The female rats which formed the parent generation were retained on the low protein diet from the age of three weeks. At 90 days of age, 10 low protein female rats and a control group of 10 female rats were mated with normal male rats from the laboratory stock. The offspring of the low protein group and of the control group were kept with their mothers until they were 35 days of age. At no time during the period of breast-feeding was the diet of the low protein mothers supplemented, though both the control and the experimental rats had unrestricted access to their respective diets.

The results reported in the study refer to the growth, mortality, intelligence, and exploratory behaviour of the first filial generation pups. Some observations made on the initial parent generation and on a second parent generation selected in order to obtain confirmation of the results, are also reported.

Retardation of growth, as reflected in weight curves, was observed in both the parent and the first filial generations. Mortality rates were high in the low protein pups of the first filial generation. Seventy per cent of the deaths occurred between the 18th and 20th day after birth. In the mature low protein rats of this generation there was also a significantly higher number of deaths than in the control group. At autopsy, lesions of the lungs were found in all of the low protein rats, and, in five of the nine, lesions of the alimentary tract were present. The specific cause of death was not established.

Minor modifications in the administration of the Hebb-Williams test of

intelligence were introduced, and the preliminary training period on this test was used to provide an index of exploratory behaviour. Neither in the parent generation, nor in the first filial generation, were there any marked differences between the low protein rats and the control rats in exploratory behaviour.

The parent generation was not tested on the problems of the Hebb-Williams test. A subsequent study has shown, however, that parent generation rats retained on a low protein diet from the age of three weeks do not score differently on the test from a comparable control group.

There were marked differences in intelligence between the low protein rats of the first filial generation and the control group. The low protein rats scored lower on the problems of the Hebb-Williams test and took longer to reach the goal box.

It is suggested that the differences observed between the parent and the first filial generations may be accounted for by the effects of the low protein diet fed to the parent generation on the pups of the first filial generation. The low protein diet of the mothers could well have affected the first filial generation during gestation. Further, the milk of the low protein mothers of the parent generation may have been defective in quality or quantity, or both, during the period of lactation. Either or both of these periods could well have been crucial in influencing the development and behaviour of the first filial generation.

The results are discussed in relation to certain dietary and feeding patterns which have been reported as being characteristic of many African peoples.

REFERENCES

1. ANDERSON, T. E., & SMITH, A. H. Effects of quantitative and qualitative stunting on maze learning in the white rat. *J. Comp. Psychol.*, 1926, **6**, 337-359.
2. AUFFRET, C., & TANGUY, F. La méthionine dans le lait des femmes africaines de Dakar. *Bull. Med. Afr. Occid. Franc.*, 1949, **6**, 99-101.
3. BERNHARDT, K. S. Protein deficiency and learning in the rat. *J. Comp. Psychol.*, 1936, **22**, 269-272.
4. BEVAN, W., LEWIS, G. T., BLOOM, W. L., & ABESS, A. T. Spontaneous activity in rats fed an amino acid deficient diet. *J. Physiol.*, 1950, **163**, 104-110.
5. BEVAN, W., & FREEMAN, O. I. Some effects of an amino acid deficiency upon the performance of albino rats in a simple maze. *J. Genet. Psychol.*, 1952, **80**, 75-82.
6. BROCK, J. F., & AUTRET, M. Kwashiorkor in Africa. Geneva: World Health Organisation, Monograph Series, 1952, **8**, 5-78.
7. DREYER, J. J. The relative biological values of the proteins of breads made from wheaten flour of 90 per cent extraction with or without supplements of "deodorised" fish flour. (Unpublished report.) Pretoria: South African Council for Scientific and Industrial Research, National Nutrition Research Institute, 1958.

8. ESPE, D., & SMITH, V. R. *Secretion of Milk*. (4th ed.) Ames: Iowa State College Press, 4th ed., 1952.
9. GILLMAN, J., & GILLMAN, T. *Perspectives in Human Malnutrition*. New York: Grune & Stratton, 1951.
10. GRIFFITHS, W. J., & SENTER, R. J. The effect of protein deficiency on maze performance of domestic Norway rats. *J. Comp. & Physiol. Psychol.*, 1954, 47, 41-43.
11. HEBB, D. O., & WILLIAMS, K. A method of rating animal intelligence. *J. Gen. Psychol.*, 1946, 34, 59-65.
12. KON, S. K., & MAWSON, E. H. Human milk, wartime studies of certain vitamins and other constituents. London: Her Majesty's Stationery Office, 1950.
13. MACHELLA, T. E., & GRIFFITH, J. Q. In Farris, E. J., & Griffith, J. Q. (Eds.), *The Rat in Laboratory Investigation*. (2nd ed.) Philadelphia: Lippincott, 1949.
14. MILLER, J., & PLATT, B. S. Chronic protein malnutrition in the rat and the effects of supplements of fish meal. (Conferência Inter-Africana de Nutrição.) Luanda: Secretariado geral da conferência, 1956.
15. PILGRIM, F. J., ZABAREKO, L. M., & PATTON, R. A. The rôle of amino acid supplementation and dietary protein level in serial learning performance of rats. *J. Comp. Psychol.*, 1951, 44, 26-35.
16. POLLACK, H., & HALPERN, S. L. In Anson, M. L., Edsall, J. J., & Bailey, K. (Eds.), *Advances in Protein Chemistry: The Relation of Protein Metabolism to Disease*. New York: Academic Press, 1951.
17. RAOULT, A. Clinical aspects of Kwashiorkor in French West Africa. In *Malnutrition in African Mothers, Infants, and Young Children*. (Report of second Inter-African conference on nutrition.) London: Her Majesty's Stationery Office, 1954.
18. RIESS, B. F., & BLOCK, R. J. The effect of amino acid deficiency on the behaviour of the white rat: Lysine and Cystine deficiency. *J. of Psychol.*, 1942, 14, 101-113.
19. SHAY, H., KOMAROV, S. A., FELS, S. S., MERANZE, D., GRUENSTEIN, M., & SIPLET, H. A simple method for the uniform production of gastric ulceration in the rat. *Gastroenterology*, 1945, 5, 43-61.
20. SLONAKER, S. R. The effect of different percentages of protein in the diet of six generations of rats. California: Stanford Univ. Publications. *Biol. Sci.*, 1939, 6, 257-321.
21. SMITH, E. A., & CONGER, R. M. Spontaneous activity in relation to diet in the albino rat. *Amer. J. Physiol.*, 1944, 142, 663-665.
22. SQUIRES, B. T. The feeding and health of African school children. (Sch. Afr. Studies.) Cape Town: University Press, 1949.
23. ———. Malnutrition in the Bechuanaland Protectorate. In *Malnutrition in African Mothers, Infants, and Young Children*. (Report of second Inter-African conference on nutrition.) London: Her Majesty's Stationery Office, 1954.
24. TROWELL, H. C., DAVIES, J. N. P., & DEAN, R. F. A. Kwashiorkor. London: Arnold, 1954.
25. WYN-HARRIS, SIR PERCY. Inaugural address, Second Inter-African conference on nutrition. In *Malnutrition in African Mothers, Infants, and Young Children*. London: Her Majesty's Stationery Office, 1954.

Developmental Research Unit
University of Natal
Pietermaritzburg, South Africa

A METHOD FOR APPRAISING CHILDREN IN IN-PATIENT PSYCHIATRIC TREATMENT: RANKING FOR NORMALITY*

Henry Ittleson Center for Child Research, New York

WILLIAM GOLDFARB

Psychiatric diagnosis of a child in accord with current nosology has limited theoretic and practical value. If, for example, a given child is classified as suffering from childhood schizophrenia, the designation merely acts as a gross alerting signal that pronounced ego defects may be anticipated. Yet the term refers to a very heterogeneous group of children. There is a distinct likelihood of multiple causes, each with different prognostic value—even in the face of symptomatic uniformity. Further, if one considers the strength, complexity, and level of integration of the self regulating, organizing functions of the child's ego, there is great diversity among schizophrenic children. An awareness of this diversity is a necessity in planning treatment and daily management of these children. Certainly the mere designation of a child as "childhood schizophrenic" does not suggest a specifically helpful life design for him nor is it a relevant basis for defining the direction and intensity of individual psychotherapy. Finally, the term "childhood schizophrenia" does not define the clinical status and progress of a child. These may vary widely within the very general bounds of the diagnosis.

To be of practical use, therefore, psychiatric classification needs to include a detailed analysis of the adaptive equipment available to the child. Frequently, in clinical practice, an assay of overall or global level of ego organization is found helpful in formulating a treatment plan for a child. This type of evaluation is of value, too, in appraising a child's progress in treatment. Ordinarily, in such an evaluation, the child is compared with the other children in treatment. He may, for example, be described as belonging to the upper, middle, or lower ranges in ego integration. In effect, this means that each child in the treatment group is ranked grossly with regard to his similarity to the presumptively normal child. Following is a report of such a procedure. At the time of this study in 1955, there were 21 children in in-patient psychiatric treatment at Ittleson Center. These children had been studied and observed professionally for periods up to two and a half years.

* Received in the Editorial Office on June 2, 1958.

They were very familiar, therefore, to all the professional personnel including psychiatrists, caseworkers, counsellors, and teachers. Each professional worker was asked to rank the 21 children in their approximation to normality. The raters were given the following directions:

You are asked to coöperate in estimating how you see Ittleson children in order of normality—which child is most normal, which is next more normal, and so on down to the child who is least normal.

Each of you has a deck of cards. On each of the cards is the name of one of the children. Go through the deck, pick out the child who impresses you as the most normal in his total adjustment and behavior. Place that card on top. Then from the remaining cards pick out the next child who is most normal. In this manner, rank all of the children from most nearly normal on top to least normal on the bottom. If two children appear equally normal, place them next to each other.

When you have finished your ranking, put your name on the top card and place a rubber band around the whole deck.

Please rank the children without assistance from anyone else and do not discuss what you have done with anyone else. The results will be discussed with you as soon as the data are in.

Each child was ranked in this fashion by 26 adults, including nine counsellors, one caseworker, one nurse, two supervising counsellors, one psychologist, four psychiatrists, four members of the research staff, and three teachers. Correlations within classes of personnel were first computed and then the average ranks given by each class of personnel were intercorrelated. For estimating the interrelations within classes of personnel, Kendall's coefficient of concordance (W) was used (1). Table 1 indicates the value of W and the equivalent average rank correlation coefficient for each class of personnel. It is clear that within each class of personnel there is a remarkable consistency in the ranking of the 21 children. None of the differences among the W 's is significant. Whether seen from the viewpoint of counsellors,

TABLE 1
THE AGREEMENT OF JUDGES WITHIN A PROFESSIONAL CLASSIFICATION: THE COEFFICIENT OF CONCORDANCE (W) AND THE EQUIVALENT AVERAGE RANK-CORRELATION

	Coun- sellor	Case- worker	Nurse	Super- vising counsellor	Psychol- ogist	Psychi- atrist	Re- search work- ers	Teacher
W	.87	.84	—*	.97	—*	.88	.92	.84
Average rank correlation Coefficient	.86	.84	—	.97	—	.85	.89	.76

* W not computed since only one person was within professional classification.

psychologist, psychiatrist, research workers, or teachers, therefore, the appraisals of the children are highly consistent within each group of personnel.

Since each group of personnel is remarkably consistent within itself in the ranking of normality of the children, the next question would be: Do the various groups of personnel agree with one another? The average intercorrelations among the several classes of professional personnel are also unusually high (Table 2).

TABLE 2

The agreement among classes of professional personnel in ranking the approximation to normality of Ittleson children: The intercorrelations among the sums of ranks for children by any one group of professional personnel with that of any other group of personnel.

Classes of personnel	Counsellor	Case-worker	Nurse	Supervising counsellor	Psychologist	Psychiatrist	Research workers	Teacher
Number in class	9	2	1	2	1	4	4	2
Counsellor	—	.88	.88	.91	.82	.92	.92	.97
Caseworker		—	.85	.93	.88	.95	.90	.91
Nurse			—	.81	.83	.84	.89	.88
Supervising counsellor				—	.82	.93	.89	.90
Psychologist					—	.89	.93	.88
Psychiatrist						—	.84	.95
Research							—	.96
Teacher								—

In view of the very high consistency in these rankings, a practical rating of normality status is an average of all 26 ranks for each child. A standard deviation of ranks for each child is also computed as a measure of uniformity of appraisal. A picture of the in-patient population of the 21 children in this study with regard to normality is shown in Table 3. These children are being ranked in this way at regular intervals. Shifts in clinical status are immediately apparent in the changes in average normality rank. Children admitted to the Ittleson Center in periods between whole group rankings receive a normality rank on the basis of a man-to-man rating using the original 21 children as referents. Each judge is asked to place the new child between the two originally rated children whom he most closely resembles in normality. Ordinarily the judges do this with little difficulty. The new child is given a rank order between the ranks assigned the two previously rated children.

The rank-order appraisal of normality level of emotionally disordered children in in-patient treatment is a simple and feasible procedure. There is

TABLE 3
THE AVERAGE RANK GIVEN EACH OF THE 21 CHILDREN TOGETHER WITH THE STANDARD
DEVIATION OF THE 26 RANKS ASSIGNED EACH CHILD

Child No.	Average rank	Standard deviation
1.	16.00	2.2
2.	17.5	1.3
3.	10.2	3.2
4.	10.2	4.0
5.	12.9	2.3
6.	3.6	2.2
7.	10.6	2.8
8.	19.5	1.1
9.	4.2	3.0
10.	3.4	1.9
11.	8.0	2.2
12.	16.5	1.5
13.	10.0	3.4
14.	7.6	3.1
15.	21.0	0.3
16.	19.0	0.7
17.	9.4	2.7
18.	9.3	3.4
19.	2.7	2.4
20.	15.2	2.0
21.	5.5	2.6

also extraordinary uniformity in this type of appraisal of the same child by a diverse group of professional observers. The method, therefore, is regarded as an excellent measure of a child's closeness to the normal child in overall ego organization and one which lends itself easily to therapeutic and investigative use.

REFERENCE

1. KENDALL, M. M. *Advanced Theory of Statistics*. New York: Lippincott, 1943.
Henry Ittleson Center for Child Research
5050 Iselin Avenue
Bronx 71, New York

PARTIAL REINFORCEMENT FOLLOWING A LONG CONTINUOUS REINFORCEMENT ACQUISITION SERIES*¹

Department of Psychology, Northwestern University

DONALD J. LEWIS AND DONALD A. HARDER

A. PROBLEM

The primary purpose of this study was to determine the effects of partial reinforcement subsequent to a long continuously reinforced acquisition series. The long continuously reinforced acquisition series preceding the onset of partial reinforcement should bring habit strength to a maximum and thus afford a means of determining whether partial reinforcement has its effect upon learning or performance. A common assumption (1, 2) is that learning is relatively permanent. Therefore, if a partially reinforced response, after having been brought to asymptotic strength, is more resistant to extinction than a continuously reinforced one, also at asymptotic strength, then the partial effect is due to some performance variable and not to superior learning.

B. METHOD

1. *Subjects*

Two randomly assigned groups, each composed of 10 male albino rats of the Sprague-Dawley strain, were used. At the beginning of the experiment Ss were approximately 90 days old with a weight range from 185 to 222 gms.

2. *Apparatus*

An L-shaped flat black maze was used, the alley of which was four ft. long and the arm was one ft. long. Both alley and arm were four in. wide, and all walls were nine in. high. The alley and arm were covered with mosquito netting and the 6 x 4 x 9 in. start and goal boxes were covered with glass. The only illumination was a 60 watt bulb suspended 6 ft. over the center of the maze. Running times were obtained to the nearest .10 sec. by means of a stop watch.

* Received in the Editorial Office on June 2, 1958.

¹ This study was supported by a grant from the National Science Foundation.

3. Procedure

All *Ss* were placed on a 22-hr. food schedule seven days before experimentation, and they were given five min. of handling and taming on each of the last five scheduling days. The reinforcements were .45 mg. dry pellets obtained from the P. J. Noyes Co. Water was available at all times in the living cages where *Ss* stayed during the intertrial interval. The general pre-experimental treatment was similar to that used previously in this laboratory (2).

All *Ss*, 10 in each of two groups, were given identical acquisition training for 100 trials. Ten trials were given a day for 10 days. Each *S* was placed in the start box and one sec. later the door was opened. As soon as he entered the alley the start box door was closed. When *S* made the turn into the arm of the *L*, another door was closed behind him. The *Ss* were retained in the goal box for 10 sec. and then returned to their living cages for the 20 sec. intertrial interval.

After the 100th trial the experimental procedure was initiated. The experimental *Ss* received 30 partially reinforced trials, all on one day. The control *Ss* received a reward on each of a comparable set of 30 trials. Rewards were removed from the goal box on the 131st trial for all *Ss* and extinction began. The *Ss* continued to run until a criterion of three min. of "no responding" was reached. If *S* had not reached the goal box three minutes after the opening of the start box, he was considered to have extinguished.

C. RESULTS

All running times were converted into running speeds—reciprocal running time—and analyses were performed on these transformed scores. Medians of five trial blocks were determined and these medians are the points graphed in Figure 1.

From Figure 1 it can be seen that both groups had about reached an asymptote by the 100th trial. There were no differences, as determined by *t* tests, between the two groups either for the mean of all 100 trials or for the mean of the last five of the 100 trials. For the next 30 trials the control group continued to receive 100 per cent reinforcement, but the partial group was switched to 50 per cent reinforcement. After the 100th trial the two running speed curves begin to diverge markedly. The divergence is almost solely the result of the slowing up of the partial group; the continuous group remained at an asymptotic level. A *t* of 3.06, significant

at better than the .01 level for 18 *df*, was obtained from an overall test of these 30 trials.

The total number of trials to the extinction criterion of 3 sec. without reaching the goal box was taken as the measure of resistance to extinction. The mean number of trials to this criterion for the control group was 21.3;

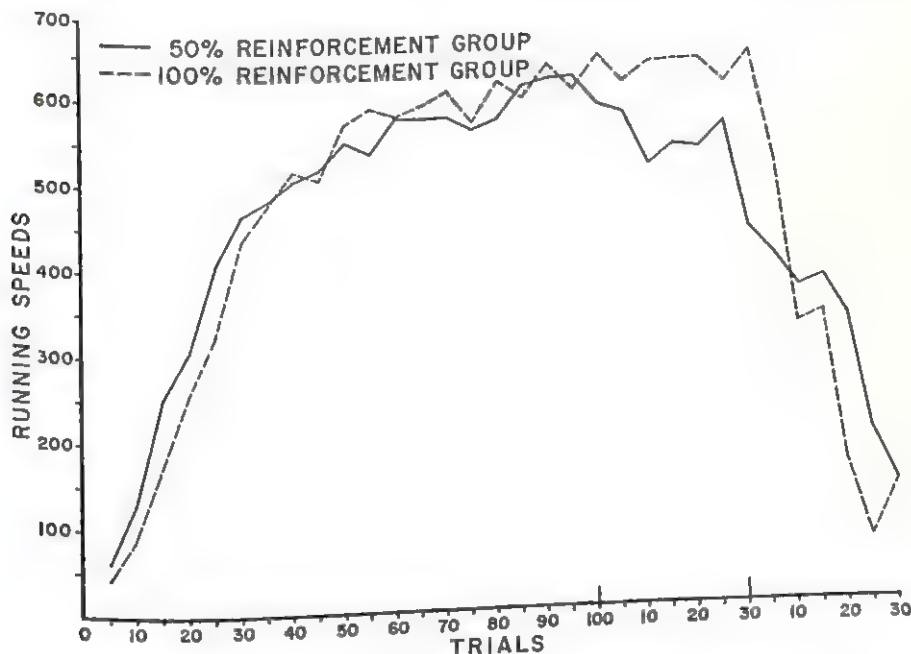


FIGURE 1

THE RELATIONSHIP BETWEEN RUNNING SPEEDS AND TRIALS

The first 100 trials were continuously reinforced for all animals. The next 30 trials were either continuously or partially reinforced. The last 30 trials were extinction.

the mean number of trials for the partial group was 28.1. This difference was also significant ($t = 3.63$) at better than the .01 level. The 50 per cent Ss thus took significantly longer to extinguish than did the 100 per cent Ss. Because some of the Ss reached the extinction criterion early, the points for the last 15 extinction trials are not based on 10 Ss. By the 25th trial only four Ss were left in the continuous group while nine still remained in the partial group. By the 30th trial all but one of the continuous group Ss had extinguished, but five animals still remained in the partial group.

D. DISCUSSION

The chief concern of the present experiment was to study the effect of partial reinforcement on acquisition and extinction after a long continuous reinforcement acquisition series. If we accept the common assumption that learning (habit strength) is relatively permanent while performance may vary, then the dip in the curve following the 100th acquisition trial is a reflection of a performance change and not a learning change. Partial reinforcement, then, clearly has its effect on performance and not on learning.

The dip in the curve following the onset of 50 per cent reinforcement can only mean that partial reinforcement has at least an initial depressant effect under the present conditions. If frustration has resulted from these nonreinforcements, it certainly did not serve to increase response strength.

E. SUMMARY

Two groups of 10 rats were given 100 continuously reinforced acquisition trials, bringing response strength virtually to an asymptote. Then one group remained on a continuously reinforced schedule for 30 trials, and the other was shifted to a partial schedule for 30 trials. Both groups were then extinguished. The results indicated that:

1. The onset of partial reinforcement had at least an initial depressant effect.
2. Partial reinforcement, even after a long continuously reinforced series, resulted in increased resistance to extinction.
3. Partial reinforcement has its effects on a performance variable and not upon increased learning.

REFERENCES

1. LEWIS, D. J., & COTTON, J. W. Learning and performance as a function of drive strength during acquisition and extinction. *J. Comp. & Physiol. Psychol.*, 1957, 50, 189-193.
2. SPENCE, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. Exp. Psychol.*, 1953, 45, 57-63.

Department of Psychology
Louisiana State University
Baton Rouge, Louisiana

Department of Psychology
University of Wisconsin
Madison 6, Wisconsin

A SCALE TO IDENTIFY IMPULSIVE BEHAVIOR IN CHILDREN*

Department of Psychology, Bowling Green State University, Ohio

B. SUTTON-SMITH AND B. G. ROSENBERG¹

A. THE PROBLEM

Impulsive behavior in children is a problem of considerable practical importance to clinicians, group workers, and teachers. However, its theoretical interrelationships with other personality variables have been only partly indicated and there is, as yet, no simple test for measuring the presence or absence of this trait (2). Those who have studied the trait systematically have not provided such tests. In *Explorations in Personality* (5), Murray includes a general personality trait called "impulsion-deliberation" and a specific sensori-motor variable called "motor-deliberation." In Murray's study, however, individuals were rated on impulsion-deliberation only after an intensive series of tests, questionnaires, and ratings, and on motor-impulsion from a computing of the errors made on a sensori-motor learning task. Sanford (7) and Polansky (6) arrived at measurements of this variable using psychologist and group leader ratings. Hartshorne and May (3) used a number of ingenious but difficult to manage party-game situations as a measure of inhibition (and by implication, impulsivity) in children.

The authors of this paper were interested in the possibility of devising a scale which would reliably and economically distinguish the most impulsive from the least impulsive in any group of children. Such a technique would further the understanding of the relationships of impulsive behavior to other personality variables, the development of normative data based on sex and age, and a more sophisticated analysis of the situational factors which tend to inhibit or release this behavior in children.

B. PROCEDURE

The first step in the present study was the creation of a pool of test items which held promise for differentiating impulsive from nonimpulsive behavior. A questionnaire containing 113 items was devised which held promise of dis-

* Received in the Editorial Office on June 2, 1958.

¹ This study was facilitated by a grant from the Scholarly Advancement Committee, Bowling Green State University.

criminating impulsivity. The bases for selecting and creating these items were adopted from Murray who defines "impulsion" as a tendency to respond quickly and without reflection, and states that it is a rather coarse variable including such elements as short reaction to social press, quick intuitive behavior, emotional drivenness, lack of forethought and readiness to work without a carefully constructed plan. In addition the investigators' experience with children suggested that impulsivity could be identified by such terms as "restlessness," "motor overflow," "horseplay," "code-breaking," and "his-trionics."

The questionnaire was administered to a group of college freshmen ($N = 86$) along with the *Ma* scale on the Minnesota Multiphasic Personality Inventory. The *Ma* scale apparently assesses impulsive behavior (4) and was therefore selected as an available criterion. Item analysis of most and least impulsive individuals (upper and lower quarters) on the *Ma* scale yielded 35 items. The first 15 items showed good differentiating power (.05 level of confidence, or better) and were retained as Impulsivity Scale One (IMP_1). The remaining 20 items showed promise (.30 level of confidence or better, but greater than .05) and were retained. A second group of college freshmen ($N = 123$) was given the 35 items along with the *Ma* scale. On the IMP_1 scale, the second group of most and least impulsive students was differentiated at better than the .01 level of confidence ($t = 7.395$, 58 *df*, $P = .01$).

These 35 items were then modified into terms more comprehensible to children. The items were administered to a group of 26 children² with the following instructions: "Read carefully each question. Put a circle around the word *Yes* if you think that it is true about you. Circle the word *No* if you think it is not true about you." Space was provided on the test sheet for name, grade, age, sex, and teacher. The first 15 items (IMP_1) which had been found to be the best differentiators of college students were found also to be the best differentiators of the children when the test scores were correlated with the teacher's independent ranking of the children on impulsivity ($r = -.53$, $P = .03$). The test-retest reliability coefficient with IMP_1 ($N = 26$) was $+.85$. Thus was a children's impulsivity scale derived indirectly from its usefulness with college students.

The authors were interested further in the possibility of deriving an impulsivity scale directly from children on the basis of teacher ratings. To this

² These subjects and others made available for the study were through the kindness of Miss E. Lorenz, Mr. N. Jacobs, Mrs. B. Sima, and Mr. R. Knestrict.

end the same 35 items were administered to three classes of children (4th, 5th, and 6th grades). On the basis of teacher ratings, the highly impulsive children ($N = 21$) were differentiated from children of low impulsivity ($N = 21$) on 10 items. Six of these items were contained in the original IMP_1 scale, and four came from the remaining 20 of 35 items. Thus there were two scales: IMP_1 containing 15 items derived from adults; and IMP_2 containing 10 items derived from children. Table 1 lists these scales and their common items.

TABLE 1
ITEMS INCLUDED IN THE IMP_1 SCALE AND RESPONSES SCORED FOR IMPULSIVITY

1. I like to keep moving around.	(Yes)
2. I like to just "blow off" steam.	(Yes)
3. I like to wrestle and to horse around.	(Yes)
4. I must admit I'm a pretty good talker.	(Yes)
5. Sometimes I can hardly stop from throwing snowballs at people I see walking by.	(Yes)
6. I don't think you should always have to do what you are told.	(Yes)
7. I like to go with lots of other kids, not just one.	(Yes)
8. It's hard to stick to the rules if you're losing the game.	(Yes)
9. I play hooky sometimes.	(Yes)
10. I don't think I'm as happy as other people.	(Yes)
11. I like to dare kids to do things.	(Yes)
12. I like throwing stones at targets.	(Yes)
13. I get into tricks at Halloween.	(Yes)
14. I'll try anything.	(Yes)
15. I am restless.	(Yes)

Items Included in the IMP_2 Scale and Responses Scored for Impulsivity

1. I like to keep moving around.	(Yes)
2. I like to wrestle and to horse around.	(Yes)
3. I must admit I'm a pretty good talker.	(Yes)
4. Sometimes I can hardly stop from throwing snowballs at people I see walking by.	(Yes)
5. I don't think I'm as happy as other people.	(Yes)
6. I like throwing stones at targets.	(Yes)
*7. I usually say the first thing that comes into my head.	(Yes)
*8. I often act on the spur of the moment without stopping to think.	(Yes)
*9. Whenever there's a fire-engine going someplace, I like to follow it.	(Yes)
*10. I like to shoot with bows and arrows.	(Yes)

* Additional items in IMP_2 .

In order to test the discriminating value of the two scales separately (IMP_1 and IMP_2) and combined (IMP_3), six additional classes were selected for study (two each of 4th, 5th, and 6th grades). IMP_3 is the result of combining IMP_1 and IMP_2 and can be seen by examining Table 1. It includes 19 items (including six common items). The students in each class were given the $IMP_{1,2,3}$ scales on a group basis with numbers assigned to each student rather than a name, in order to reduce any tendency to

falsify answers. The classroom teachers ranked the students in their respective classrooms on impulsivity, 1 being the rank assigned to the most impulsive child, and so on. Thus, each child received one ranking by his classroom teacher (all incomplete protocol and protocol of nonreaders being eliminated). Results of correlating teacher rankings with the impulsivity scales are found in Table 2.

TABLE 2
CORRELATION COEFFICIENTS AND TESTS OF SIGNIFICANCE OF IMPULSIVITY SCALES AND
TEACHER RANKINGS

Class	N	IMP ₁ (Adult derived)		IMP ₂ (Child derived)		IMP ₃ (Combined)	
		r	P	r	P	r	P
4	31	-.62	<.01	-.62	<.01	-.46	.01
4	29	-.40	.03	-.39	.04	-.54	<.01
5	30	-.40	.03	-.27	>.05	-.33	>.05
5	27	-.24	>.05	-.33	.08	-.39	.06
6	26	-.64	<.01	-.73	<.01	-.47	.02
6	28	-.42	.02	-.37	.05	-.41	.03
N = 171							

As can be seen from Table 2, each scale correlates significantly with the criterion rankings except in a single case.

Table 3 presents by grade and sex the descriptive statistics for IMP₃ based on the data obtained from 334 children. A general tendency for boys to receive higher scores than girls can be noted ($t = 8.22$, $P = <.01$). Grade does not appear to be a differentiating factor, though the difference between fourth and fifth grades is significant at the .01 level. The authors are unable to explain this difference at present. Plots of the distributions on IMP₃ by sex reveal highly comparable distributions with a slightly more positive skew for the girls.

TABLE 3
IMPULSIVITY SCALE (IMP₃) MEANS AND SD'S FOR DIFFERENT GRADE LEVELS AND FOR
BOYS AND GIRLS SEPARATELY

	Fourth Grade			Fifth Grade			Sixth Grade		
	N	M	SD	N	M	SD	N	M	SD
Boys	54	8.70	3.38	61	9.26	3.54	56	8.66	3.34
Girls	56	4.89	2.69	51	6.86	3.52	56	6.11	2.84
	110			112			112		

C. DISCUSSION

While there would be some value in presenting only a single scale (IMP_3) resulting from the two approaches to the measure of impulsivity, the authors have chosen to present the two original scales (IMP_1 and IMP_2) because such presentation has the incidental advantage in demonstrating that one can derive substantially the same scale from utilizing two distinct approaches to the same problem. If anything, it suggests corroboration of the goodness of those items which are common to both scales. It also provides a check on the validity of the method of deriving a scale for use with children which has been constructed with adults (see Castaneda *et al.*, 1).

Three correlation coefficients relating teacher rankings to IMP_1 , IMP_2 , and IMP_3 (see Table 2) fail to achieve significance. It is possible that increasing the size of the respective samples would have yielded coefficients that attain significance.

Grade appears to be a factor in determining performance on the Impulsivity Scale in only one instance. Whether or not this can be related to chance factors is unknown with the present sample.

The authors suggest that the combined scale (IMP_3) is a useful and economical instrument for isolating extremes of the behavior concerned.

D. SUMMARY

The present investigation sought to derive a scale which would reliably distinguish the most impulsive from the least impulsive children. Two approaches to scale derivation, based on adult responses and one based solely on the responses of children furnished two scales. When combined, these yielded a third scale containing 19 items. This scale was validated by using teacher ratings with a sample of 171 children. Test-retest reliability was $+.85$. Boys were found to score significantly higher than girls on the impulsivity scale. Grade was not uniformly influential in determining performance on the scale.

REFERENCES

1. CASTANEDA, A., MCCANDLESS, B. R., & PALERMO, D. The children's form of the Manifest Anxiety Scale. *Child. Devel.*, 1956, 27, 317-326.
2. CARMICHAEL, L. Manual of Child Psychology. (2nd. ed.) New York: Wiley, 1954.
3. HARTSHORNE, H., MAY, M. A., & MALLER, J. B. Studies in Service and Self-control. New York: MacMillan, 1929.
4. HATHAWAY, S. R., & MCKINLEY, J. C. The Minnesota Multiphasic Personality Inventory. Minneapolis: Univ. Minnesota Press, 1943.
5. MURRAY, H. A. Explorations in Personality. New York: Oxford Univ. Press, 1938.

6. POLANSKY, N., LIPPITT, R., & REDL, F. An investigation of behavioral contagion in groups. *Hum. Rel.*, 1950, **3**, 319-348.
7. SANFORD, R. N. Personality patterns in school children. In Barker, R. G., Kounin, J. S., & Wright, H. F. (Eds.), *Child Behavior and Development*. New York: McGraw-Hill, 1943.

Department of Psychology
Bowling Green State University
Bowling Green, Ohio

THE EFFECT OF STIMULANT AND DEPRESSANT DRUGS ON A MEASURE OF EMOTIONAL REACTIVITY IN THE RAT*

Institute of Psychiatry, Maudsley Hospital, University of London

P. L. BROADHURST, S. N. SINHA, AND S. D. SINGH¹

A. INTRODUCTION

Hall's open-field test (12) has been widely used as a measure of emotionality in the rat. It is based on the tendency of the animal to display emotional elimination (defecation and urination) and exploratory behavior when placed in a strange, fear-provoking environment, the open field. The validity of the test is reasonably well-established: some of the difficulties associated with its use which have led to doubts on this score have been overcome by standardising the test conditions at a slightly more stressful level than that previously used (see 3).

The purpose of the present experiment is to explore the effect upon performance in this test of various drugs. In order to maximise the possibility of encountering a measurable effect which might throw light on the psychological processes involved, both stimulant and depressant drugs were used, and ones were sought which are known principally for their action on either the central or the autonomic nervous system. While the overall stimulant or depressant action of a drug is usually reasonably well established, the level or site of such action is frequently obscure. Consequently, it will be appreciated that, in the present state of pharmacological knowledge, it is not possible to select drugs which conform exactly to the fourfold classification (of stimulant/depressant, with central/autonomic action) as described. This classification, moreover, is itself doubtless an oversimplification. It provides, nevertheless, a rational basis for selecting the drugs according to their principal properties.

The drugs used were pipradrol hydrochloride ("Meratran," Rikers), a

* Received in the Editorial Office on June 6, 1958.

¹ We are indebted to Rikers Laboratories, Loughborough, Leics., England, for the supply of "Meratran" used in the first part of the experiment; to Drs. S. E. Dicker and R. H. J. Watson of University College, London, for suggesting the trial of methyl-pentynol and assisting with the supply of drugs used in the second part; and to Dr. Watson, and also to Mr. A. F. Green of The Wellcome Research Laboratories, for their valuable criticisms.

central stimulant (8); amobarbital sodium ("Sodium Amytal," Lilly), a depressant; ephedrine hydrochloride, which has a peripheral action but acts as an autonomic stimulant; and methylpentynol, ("Oblivon," British Schering), which decreases autonomic reactivity (9).

B. METHOD

The experimental design involved administering the open-field test twice to experimental and to control groups. On the first occasion no drugs were used, whereas the second or re-test was given under the influence of the appropriate drugs or placebos respectively. The effect of any drug action will thus be seen by comparing the difference between the test and the re-test scores of experimental animals with the difference between those of control animals. This procedure, which is illustrated in Table 1, avoids difficulties

TABLE 1
EXPERIMENTAL DESIGN

	Open-Field Test	
	First	Second
Experimental groups	No drug	Drug
Control groups	No drug	Placebo

in equating groups with respect to their defecation scores, since it is possible by this method to use the scores derived from the first test in equating the groups given the various treatments for the second test.

1. Open-Field Test

The open-field test was administered on both occasions in the revised and standardised version, the development of which is described elsewhere (3). Essentially it consists in exposing the subject in a uniform manner for a two-minute period on four successive days within a circular arena, 32 3/4 ins. in diameter, with white-painted plywood walls 12 1/2 ins. high, to constant sound ("white" noise) and light fields whose intensity at floor level averaged 78 db. (ref. 0.0002 dynes/sq. cm.) and 165 fc. respectively. Above and around the arena a light metal structure supports the batteries of loudspeakers and lamps for providing the noise and light stimulation as well as a screen of muslin which forms a rough one-way screen. The noise also serves as a "sound screen." The subject under test is thus largely insulated from external distraction. The number of fecal boluses the rat deposits, and the number of compartments, marked on the floor, and roughly equal in size,

through which it passes, are noted, and constitute the defecation and ambulation scores respectively.

2. Subjects

All the 54 subjects used in this experiment were drawn from the fifth and sixth generations of rats selectively bred for their high or low defecation scores in the open-field test (the "Maudsley Reactive and Non-reactive strains") (7). They were first tested at 96.7 days (± 0.49 SE) of age, before which they had been subjected to strictly controlled environmental conditions not only with respect to their physical surroundings and husbandry but also to other aspects of their early environment which might affect adult behavior (1). Thus, in addition to the standardisation of diet (MRC No. 41, and tap water), air temperature (20.0 - 23.3°C), light/dark cycle (17½/6½ hrs. respectively), such variables as cage size (9 x 9 x 8½ ins. high), number of animals per cage (variable to weaning, five from 21 to 50 days, and three from 50 days to the first test around 100 days), and amount of handling given (five times between birth and first test) were strictly controlled. The subjects were assigned to the various groups in this experiment by virtue of their defecation scores in their first open-field test.

3. Experimental Design and Procedure

The experiment was conducted in two parts which differed somewhat in the techniques used, and the variables involved, especially those relating to drugs. In the first part the effects of the drugs acting mainly at the central level were studied; in the second, those of the ones acting mainly at the autonomic level.

a. *Part I (Pipradrol and amobarbital)*. Twenty-four subjects, all from the fifth generation of the selective breeding experiment mentioned above, were assigned to this part of the experiment. In addition to the variable of the three drug treatments (stimulant, depressant, and placebo), there were three other variables, each at two levels, included in the 3 x 2 x 2 x 2 factorial design. These were sex (males and females), emotional reactivity (low vs. high)—only rats depositing no fecal boluses on any day in the first open-field test were assigned to the former, and only those scoring 2.3 per day or more to the latter—and finally, previous experience. This last variable was included as a check since 12 animals had served as subjects in another experiment in which they had been trained to run an alley maze for a food reward under the influence of the same drugs in the same dosages as used here (17). Replicating our design with another group

of 12 rats which had not been used in any experimentation between the first and second open-field tests, allows any effect of this factor to be evaluated. The assignment of the groups of animals to drug treatments was done randomly.

The subjects were aged 232.5 days (± 1.02) when the re-test began. Precisely 35 mins. before each of the four daily exposures in the open-field test, each rat was injected intraperitoneally with 1 cc. distilled water/100 gm. body weight containing either pipradrol (dosage 0.1 mgm./Kg.), or amylobarbitone (0.15 mgm./Kg.), or nothing (placebo). The same dosages as those used in the related experiment mentioned above were employed in order that their apparent lack of effect on emotional reactivity—in marked contrast to their effect on a measure of reactive inhibition—could be directly checked by open-field test measures.

b. Part II (Ephedrine and methylpentynol). Thirty subjects, all from the sixth generation of selective breeding, were assigned to this part of the experiment. Only one variable other than that of drug treatment was included, all the subjects being male and experimentally naïve, apart from the initial open-field test. The variable in question was the strain and, consequently, emotional reactivity test scores of these subjects' parents, which were 4.29 mean number of fecal boluses per day for a high-scoring group, and 0.0 per day for a low-scoring group. Despite this difference in parentage, the subjects themselves, however, had all scored 2.0 boluses per day or more on the first open-field test, the group mean being 2.93 per day. While this mean score is lower by 1.04 than the mean score for the rest of the males of the high-scoring strain in the sixth generation, it is higher than the mean score of the rest of the sixth generation males of the low-scoring strain by as much as 2.55. That is to say, half our subjects are fairly typical representatives of our emotionally reactive strain, whereas the other 15 come from that portion of the distribution of the nonreactive strain whose scores, because of incomplete selection as yet, still overlap that of typical reactives. Thus, we have two groups, genetically different, but phenotypically alike on their first test, and the aim here was to contrast the performance of these two groups on their re-test in order to heighten any possible effect of drug action.

The subjects were aged 109.1 days (± 1.03) when the re-test began. In this part of the experiment the drugs were injected into each animal's throat by means of a short metal cannula, four hours after food had been removed from the cages. To accustom them to this procedure, 0.2 cc./100 gm. body

weight of the blank "Oblivon" elixir were administered on the two days prior to re-testing, and, in the case of the control (placebo) group, on the four succeeding days precisely 15 min. before exposure in the open field. On these four days the elixir given the other groups now contained either methylpentynol (dosage: 125 mg./Kg.) or ephedrine (100 mg./Kg.). After the first day, the elixir used with the last group was diluted (50 per cent volume) with syrup simplex (*BP*) in order to increase acceptance by masking further the bitter taste of the ephedrine. The dosage of methylpentynol chosen was that which has been shown to increase the activity of rats significantly (10); and that of ephedrine the maximum which did not produce marked hyperactivity in most rats. In this part of the experiment, the "blind" method was used, that is to say, the experimenter doing the open-field testing did not know which drugs had been administered to the subjects. This additional precaution is not considered essential in view of the almost complete objectivity of the scores derived from this test.

C. RESULTS

In each part of the experiment difference scores were calculated by subtracting from the total defecation scores for each rat on the second test that on the first.² After a suitable constant value had been added to each score so obtained, in order to rid the data of negative values occurring when a higher score was observed on the first test than that on the second, an analysis of variance was performed.³ The ambulation scores were treated in precisely the same manner. The findings are presented in Tables 2 and 3.

TABLE 2
SUMMARY RESULTS OF ANALYSES OF VARIANCE: "CENTRAL DRUGS"

Source of variance	df	Defecation scores Variance estimate	F	Ambulation scores Variance estimate	F
Drugs	2	33.04	1.4	155,459.04	49.9***
Sex	1	2.04	0.08	31,248.17	10.0**
Strain	1	84.37	3.5	37.50	0.01
Experience	1	5.04	0.2	294.00	0.10
Residual*	18	23.74		3,114.75	
Total	23				

* Combining all interactions, none of which is significant in either case.

** Significant at the 1% level.

*** Significant at the 0.1% level.

² One rat in the group given ephedrine died after the second day of the re-test (apparently through cardiac failure). The scores obtained on the first two days of the re-test were accordingly compared with those obtained on the first two days only of the first test.

³ A significant degree of inter-group inhomogeneity of variance was observed in certain cases, but ignored in view of the evidence now available (14) that moderate departures from the assumptions underlying the analysis of variance are not crucial.

TABLE 3
SUMMARY RESULTS OF ANALYSES OF VARIANCE: "AUTONOMIC DRUGS"

Source of variance	df	Defecation scores		Ambulation scores	
		Variance estimate	F	Variance estimate	F
Drugs	2	1.25	0.04	2,078.80	1.2
Parents' strain	1	790.53	26.3***	896.53	0.5
Interaction	2	57.62	1.9	849.73	0.5
Residual*	24	30.10		1,787.88	
Total	29				

* Individuals' error variance.

*** Significant at the 0.1% level.

It will be seen that the only significant effect of drug action which has been detected is that of the central acting drugs upon ambulation in the open-field test. Analysing this effect further, we find that it is due solely to the action of pipradrol hydrochloride, which appears to increase open-field activity. Thus, the relevant mean difference scores, converted in this case to metres run per day, show an increase of 9.46 for the group injected with pipradrol, which is a significantly greater change (at the 0.1 per cent level by 2 tail *t* test) than the decrease of 2.91 for the placebo group, or that of 4.15 for the amobarbital sodium group, and these two latter do not differ significantly between themselves.

The significant sex difference which also appears in this analysis (it will be recalled that only males were used in the second part of the experiment) is a consequence of an increase in ambulation among females (mean: 2.75 m. per day) compared with a decrease among males (mean: 1.15). No significant effect due to previous testing in a different situation is observed (18).

The expression of strain differences is confined to defecation scores, and is seen most clearly in the second part of the experiment where subjects derived from the emotionally reactive strain showed an increase between test and re-test of 0.2 fecal boluses per day compared with a decrease of 2.7 for subjects from the nonreactive strain, and this difference is highly significant, as indicated in the analysis of variance (Table 3). Similarly, in the first part of the experiment, the difference between the mean decrease for the reactive rats (0.48 boluses per day) and the increase for the nonreactive ones (0.46 per day) may be regarded as bordering significance, an *F* ratio of 4.4 being required to meet the 5 per cent criterion.

D. DISCUSSION AND CONCLUSIONS

The main finding of the present experiment is that the defecation response in the open-field test appears unaffected to any significant degree by the drugs given and in the dosage levels used. This confirms the absence of differential reaction to the centrally acting drugs by the emotionally reactive and nonreactive groups in the previous experiment on response alternation in which these drugs were used (17). It must of course be noted that other doses might have given different results, but it is also possible that the defecation response itself is one which is not especially sensitive to the action of drugs. The significant effect found on response alternation in the work mentioned above supports the latter view, as does further work in progress by one of us (S.D.S.) in which large effects have been demonstrated by the use of a similar pattern of stimulant and depressant drugs in comparable doses, but this time on conditioned emotional responses (*CER*) (2).

The distinction implied here, between two different kinds of responses to fear, one a learned response (*CER*) and the other (open-field test defecation) definitely largely innate (7), may be crucial in that the learning process is probably more susceptible to the action of the same dose of a drug than responses which, while undoubtedly mediated through the central nervous system, involve inherited nervous pathways. This view corresponds, of course, to the well-known tendency of drugs to affect learned or "higher" functions before more primitive ones. On the other hand, the simple dichotomy noted is probably an over simplification, for there is no lack of contrary instances, that is, of drug action on apparently innate responses, e.g., Schwartzbaum (16) showed the effect of reserpine on nest building in rats—though here there may be a reflex response to a physiological change in body temperature induced by the drug; and Hess (13) has shown the effect of meprobamate on imprinting in ducks—though here again an element of learning is certainly involved.

Our experience thus agrees with that of Ryall (15) in that the ambulation measure in the open-field test shows a greater response to drug action. It is however only the central stimulant, pipradrol, which has a significant effect. This may be due to an increase in the speed of responding to the stimulation involved in the open-field test, as it is known that this drug significantly increases the speed of maze running (17). The significant increase in activity measures due to methylpentynol noted by Dicker, Steinberg, and Watson (10) was not observed in the present experiment. They interpret their findings in terms of a facilitation of exploratory behavior by

the action of the drug in suppressing otherwise competing fear responses, and it might have been expected that this effect would be seen in the present experiment. This is especially so because it has been clearly shown (5, 18) that the ambulation score derived from the open-field test is not affected by previous exposure to the test. That is to say, there is no adaptation to the re-test carried over from the first test and consequently the exploratory behavior is evoked again. Further work is needed to resolve this discrepancy regarding the effects of methylpentynol.

The significant strain and sex differences observed are of interest and parallel differences which are found in absolute scores on the open-field test (as opposed to the change in such scores on re-test which constitutes the measures used here). This strain difference in behavioral reaction to re-test is more marked among the subjects of the second part of the experiment despite their lack of differentiation in respect of defecation on the first open-field test given them. They were, however, drawn from a later generation than those used in the first part of the experiment and were also re-tested much sooner, and so the marked adaptation to a second test found with the defecation response (5) was consequently less attenuated by the passage of time. In addition, the animals of the low-scoring strain only had been given a fifth day of testing in the open-field, immediately following the four days of the first test, and under an increased noise stimulation. This variation in standard procedure, made for a purpose unconnected with the present experiment and otherwise irrelevant to it, may nevertheless have accelerated among the nonreactive animals the adaptation referred to above and consequently enhanced the inter-strain differences detected. However, this difference provides further evidence of the efficacy of the selection pressure maintained in establishing these high- and low-defecating strains, in addition to that already provided by the demonstration of significant differences between them in response to certain aversive drive stimuli as seen in speed of swimming (4, 6) and in the acquisition of *CER*'s (SDS, work in progress), the high defecators being the more responsive emotionally in all cases. Similarly, with respect to physiological responses, males of the two strains are observed to vary significantly in their sensitivity to drugs as measured by changes in the *CER* developed in the course of the work mentioned above, and this inter- and Watson (11) whose analysis of it suggests that it may be partly due to differences in thyroid output.

E. SUMMARY

The effect of four drugs, pipradrol hydrochloride, amobarbital sodium, ephedrine hydrochloride, and methylpentynol, upon the defecation and ambulation responses in the open-field test of rats from emotionally reactive and nonreactive selectively bred strains was observed by administering the test twice, once under standard conditions and the second time under the action of either a drug or an appropriate placebo.

Analysis of the data shows that the only significant effect of a drug was that of the central stimulant, pipradrol, which increased ambulation. Otherwise only strain and sex differences were detected. The findings are discussed in relation to other relevant work.

REFERENCES

1. BEACH, F. A., & JAYNES, J. Effects of early experience upon the behavior of animals. *Psychol. Bull.*, 1954, **51**, 239-263.
2. BRADY, J. V., & HUNT, H. F. An experimental approach to the analysis of emotional behavior. *J. of Psychol.*, 1955, **40**, 313-324.
3. BROADHURST, P. L. Determinants of emotionality in the rat: I. Situational factors. *Brit. J. Psychol.*, 1957, **48**, 1-12.
4. ———. Emotionality and the Yerkes-Dodson Law. *J. Exper. Psychol.*, 1957, **54**, 345-353.
5. ———. Determinants of emotionality in the rat: II. Antecedent factors. *Brit. J. Psychol.*, 1958, **49**, 12-20.
6. ———. A "Crespi effect" in the analysis of emotionality as a drive in rats. *Brit. J. Psychol.*, 1958, **49**, 56-58.
7. ———. Studies in psychogenetics: The quantitative inheritance of behaviour in rats investigated by selective and cross-breeding. *Bull. Brit. Psychol. Soc.*, 1958, **34**, 2A (abstract).
8. BROWN, B. B., & WERNER, H. W. Pharmacologic studies on a new central stimulant (2-Piperidyl) benzydrol hydrochloride (MRD-108). *J. Pharmacol. & Exper. Therap.*, 1954, **10**, 180-187.
9. DICKER, S. E., & STEINBERG, H. The effect of methylpentynol in man. *Brit. J. Pharmacol. & Chemotherap.*, 1957, **12**, 479-483.
10. DICKER, S. E., STEINBERG, H., & WATSON, R. H. J. Effect of methylpentynol on the activity of rats. *J. Physiol.*, 1957, **137**, 88-89P (abstract).
11. DICKER, S. E., & WATSON, R. H. J. Personal communication, 1957.
12. HALL, C. S. Emotional behavior in the rat: I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. Psychol.*, 1934, **18**, 385-403.
13. HESS, E. H. Effects of meprobamate on imprinting in waterfowl. *Ann. N.Y. Acad. Sci.*, 1957, **67**, 724-733.
14. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Cambridge, Mass.: Houghton Mifflin, 1953.

15. RYALL, R. W. Effect of drugs on emotional behaviour in rats. *Nature (Lond.)*, 1958, **182**, 1606-1607.
16. SCHWARTZBAUM, J. S. Effects of reserpine and chlorpromazine on nest building and thermal behavior in rats. *Proc. Soc. Exper. Biol.*, 1955, **90**, 275-277.
17. SINHA, S. N., FRANKS, C. M., & BROADHURST, P. L. The effect of a stimulant and a depressant drug on a measure of reactive inhibition. *J. Exper. Psychol.*, 1958, **56**, 349-354.
18. WATSON, R. H. J. The behaviour of the rat in conflict situations presented serially. Unpublished Ph.D. Thesis, University of London, 1954.

Institute of Psychiatry
Maudsley Hospital
University of London
Denmark Hill
London, S.E.5, England

THE EXPRESSION OF GUILT IN FANTASY AND REALITY*

Department of Psychology, Stanford University

KENNETH R. WURTZ¹

A. INTRODUCTION AND STATEMENT OF THE PROBLEM

The current view on socialization assumes that most of the child's social learning following the age of three or four is influenced by the outcome of the identification process. The theory of identification maintains that the child internalizes the standards of his parents and then attempts to make his behavior coincide with these internalized standards. When there is a discrepancy between the child's behavior and his internalized standards, the child experiences guilt, which is self punishment for the failure to abide by his standards of behavior. The strongly identified child should experience more guilt under these conditions than the child who has not identified.

The above discussion implies that a different model is used by the child to govern his behavior following identification than was used before. The children who have established a firm identification use a "guilt" model and behave as if the standards of behavior are within themselves. The other children appear to use a "fear" model; i.e., they act as if their behavior is based more on the rewards and punishments involved in the immediate situation, that is, the standards of behavior are external to themselves. Because of the importance of guilt in the socialization process it is imperative that some adequate measures of guilt be developed. The present study was an attempt to deal with the problem of the measurement of guilt in children.

The measurement of a child's attitudes and emotional feelings presents difficulties since the child cannot verbalize them as easily as an adult. In some respects, play is the most natural medium in which children can express their attitudes and feelings. This explains the extensive use of doll play as an experimental technique in research with children (1, 2, 3, 4, 5, 7, 9, 10, 11, 13, 14, 15, 17). In recent years, both Korner (8) and Stolz (19) extended the doll play technique with the incomplete stories technique.

* Received in the Editorial Office on June 16, 1958.

¹ This paper is based on a dissertation submitted to Stanford University in fulfillment of a requirement for the Doctorate in Philosophy. The writer wishes to thank Drs. Robert R. Sears, Lois M. Stolz, and Albert Bandura for their constructive criticism of the plans and presentation of this study.

Finally, Sears, Whiting, Maccoby and others (16) attempted a direct application of this technique to the study of guilt and identification. They developed and used eight incomplete doll play stories dealing with various forms of deviation and implicitly assumed that the child's responses were indicative of his reactions in real life situations. Heinicke (6) attempted with some success to relate these data to the child rearing practices of parents.

The present study was designed to investigate a number of implications involved in this technique for the measurement of guilt. A standard guilt-arousing situation was developed which is simple enough to be used in studies of identification variables. An experimental group and a control group of nursery school children were given a standard doll play session to establish rapport. Several days later the subjects were administered a pre-test and a post-test consisting of six matched pairs of incomplete doll play stories; an attempt was made to control the relevant variables of these stories. These stories involved deviations on the part of the child doll similar in sex to the subject. Following the pre-test the experimental group was subjected to a guilt-arousing experience and the control group was not. The guilt-arousing experience was the collapsing of a doll house which the child felt responsible for. Three of the matched pairs of stories were similar to the guilt-arousing situation and three were different from it. The completions of the stories and the immediate responses to the collapsing of the doll house were categorized and recorded.

The specific hypotheses tested were as follows:

1. There will be an increase in guilt responses to the deviation stories following the guilt-arousing situation.
2. This increase will be greatest on the stories which are similar to the guilt-arousing situation and least on the stories which are dissimilar.
3. The child's responses to the stories will be positively related to his overt behavior in the guilt-arousing situation.

B. PROCEDURE

The subjects were 38 Stanford Village Nursery School children between the ages of 3 years, 10 months and 5 years, 5 months. The 19 experimental subjects were matched on age and sex with the 19 subjects in the control group. Both groups contained 11 males and 8 females.

The subjects in both groups were given one 20-minute session of doll play in order to establish rapport and insure that the child would manipulate the dolls in the deviation stories. A complete description of the standard doll

play procedure followed may be found in Sears (15). All of this experiment, including these doll play sessions, took place in the experimental rooms of the Stanford Village Nursery School. Figure 1 is an outline of the floor plan of these rooms. The doll play sessions took place in Area II on this diagram. Following the standard doll play there was a time lapse varying from one to 19 days, and then the experiment proper took place.

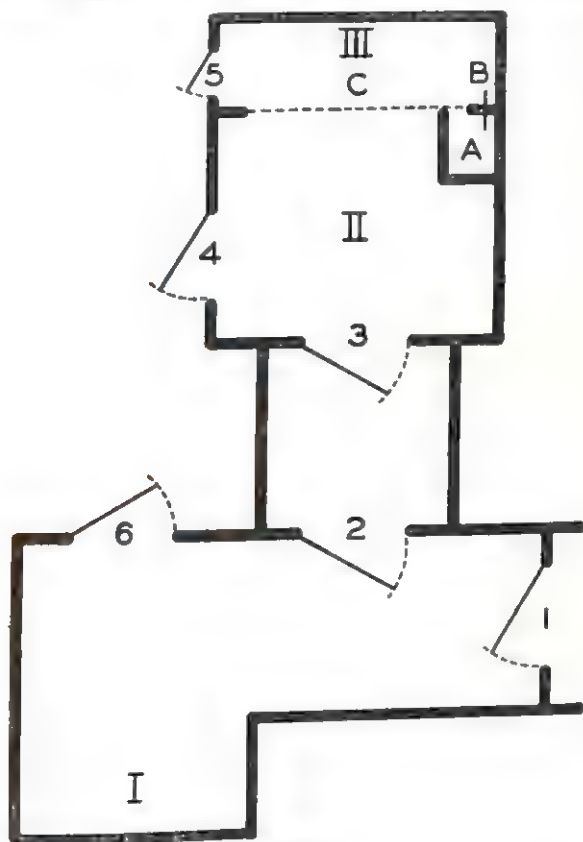


FIGURE 1
FLOOR PLAN OF THE EXPERIMENTAL ROOMS

At the beginning of the experiment, each subject was brought through Door 1 from the main section of the nursery school to Area I and given the following instructions:

Now, (name of subject), I'll tell you how we play this game. I'm going to tell you the beginning of a story, and then you will tell me how it ends. You finish the story for me. All of the children tell different endings to the stories, so you may end it any way you like.

The examiner then administered six incomplete stories to the child, using the appropriate dolls to act out each story while telling it. The dolls were given to the child and he (or she) was told to use them to finish the story. The only other props were three wooden blocks that the child could use as furniture or other objects involved in the story. Of the 12 stories used in the pre-test and the post-test, 6 were similar to the guilt-arousing situation (3 in the pre-test and 3 in the post-test) and 6 were different from it in relevant aspects (3 in the pre-test and 3 in the post-test). The similar stories involved the destruction of community property, and the dissimilar stories involved the failure to carry out a given responsibility.

The examiner observed the child's response to each story and immediately scored it in terms of the following categories:

1. Fixing—any attempt to repair damage caused by the deviation.
2. Physical Punishment—the introduction of physical pain.
3. Verbal Punishment—punishment of a verbal type.
4. Other Punishment—isolation or deprivation of privileges.
5. Feelings Punishment—the display of hurt feelings.
6. Total Punishment—the sum of the four punishment categories.
7. Confession—informing authority figure of guilt.
8. Hiding—hides either himself or the signs of the deviation.
9. Redefinition Plus—alters story so deviation is increased.
10. Redefinition Minus—alters so deviation is increased.
11. Delay—attempt to put off completing the story.
12. Subterfuge—denial of guilt by the guilty party.
13. Refusal—refusing to complete the story.

The examiner tabulated the doll agent and object for all four punishment categories and for fixing. The child was allowed to give more than one ending to each story, and the examiner waited until the child was completely finished. The next step in the procedure was the guilt-arousing experience.

The apparatus used in arousing guilt was a two-story collapsible doll house. The second floor of the house was supported by a dowel which passed through a hole in the right wall. The house was placed on a small table (*A* on the floor plan) next to the wall in Area II, and the dowel was inserted through a hole in the wall of the experimental room. The person in the observation room (Area III) was able to pull the dowel out so that the second floor of the doll house would collapse. The house contained a standard doll family and furniture.

Immediately following the pre-test, the examiner turned the child over to the experimenter and the child was taken through Doors 2 and 3 into the play room (Area II) containing the doll house. The child was told to play

with the doll house, but to be very careful not to break it. After the child played for a few seconds, the experimenter instructed the child as follows:

I have to go outside here to talk to a person for a few minutes. I'll be back in a few minutes. If you need me I'll be right outside this door. I'll be back in a few minutes.

The experimenter then left the room through Door 4, walked to Door 6, and took a position where he could view both Doors 2 and 4. From this point on the child was alone in the room with the doors closed. An observer watched the child through the one-way vision screen from the observation room. The control group subjects were allowed to play with the doll house for four minutes, and then the examiner and the experimenter came to Door 4, called the child from the room, and took him back to Area 1 for the post-test. The experimental subjects were given the guilt-arousing experience. They were allowed to play with the doll house for two minutes, and then, as soon as both of the child's hands were occupied in the second story of the doll house, the observer pulled the dowel out causing the second floor to collapse. The child was then allowed to remain in the room for two more minutes, and the observer scored the child's behavior in the following categories:

1. Fixing Plus—any realistic attempt to repair the damage.
2. Fixing minus—perseveration in unsuccessful fixing.
3. Crying.
4. Hiding—any attempt to disguise the damage.
5. Freezing—immobility.
6. Indecision and Tension—moving from one door to another and appearing upset.
7. Ignoring deviation—moves furniture to floor and plays there.
8. Flight—leaving the room through door 3 since this was an attempt to avoid experimenter and return to nursery school.
9. Confession—leaving the door through room 4 since this was an attempt to contact the experimenter.

If the child left the room during the two-minute period through either Door 3 or Door 4, he was met by both the experimenter and the examiner. The child was allowed to say whatever he wished to the experimenter, and then the examiner took him back to Area I for the post-test. If the subject did not leave the room during this period he was called from the room in the same way as a control group subject.

The administration of the stories during the post-test was the same as during the pre-test. Following the post-test, the experimental subjects were brought back to the play room and shown that the doll house was completely

repaired. The experimenter informed the child that the house really didn't break and that it was not the child's fault. This period was used to give the child support and to alleviate any possible after-effects of the guilt-arousing experience. The child was then returned to the main part of the nursery school.

C. RESULTS

Inter-scorer agreement on the story completion categories was tested prior to the study by having two examiners score the story completions of five children. Out of a total of 55 responses, the examiners agreed on 53. The child could make more than one response to each story, and the two disagreements occurred when one experimenter scored more than one response to a story and the other did not. Therefore, inter-scorer reliability is very high.

The influence of the guilt arousing situation on the story completions will be considered first. The hypothesis was that there would be an increase in the child's guilt responses to the stories following the guilt-arousing situation. Three of the experimental children refused to complete any of the stories in the post-test, so they were discarded from these comparisons. This left a total of 16 matched pairs of subjects for all comparisons except for the non-thematic categories of refusal and total change of response.

Difference scores were computed for all subjects on all of the individual categories of story responses by subtracting the pre-test score in a particular response category from the post-test score. Three such difference scores were computed for each category; one for the set of similar stories, one for the dissimilar stories, and one for all stories combined (total stories). In addition to the original scoring categories, difference scores were computed for three complex scores made up on an a priori basis: a guilt score, a fear score, and a change score.

Guilt Score: This score consists in the sum of all those categories which presumably implied a guilt reaction. These are as follows: fixing by *Cs* (child doll similar in sex to the subject) or *Un* (unspecified agent); physical, verbal or other punishment by *Un*; feeling punishment by *Cs* with *Cs* as object (crying); confession; and redefinition minus.

Fear Score: This was tabulated by summing all those categories which presumably implied a fear reaction. These are as follows: fixing by *A* (authority figure); physical, verbal, or other punishment by *Cs* (aggression); hiding; redefinition plus; and subterfuge.

Change Score: This consists in the sum of all of the difference scores,

disregarding sign, for all of the individual categories; it is an index of the number of changes that the subject made in his responses.

The significance of the differences between the difference scores of the experimental and control subjects was then tested by the *t* test for matched differences (12) and the randomization test for matched pairs (18). The *t* test could only be used with the Guilt score, the Fear score, and the Change score, since the distributions on the other scores were too skewed and contained too many ties. The only differences approaching significance were found in the Refusal category and the Change score. In the Refusal category for dissimilar stories, the experimental group shows an increase relative to the control group which is significant at the .06 level. For all stories combined this difference was significant at the .11 level. (In all cases, levels of significance refer to two-tail tests.) The experimental group showed significantly more change in their responses to dissimilar stories than the control group; this difference was significant at the .01 level. For all stories combined this difference was significant at the .08 level.

One interpretation of this lack of change from the pre-test to the post-test is that it is due to a confabulation of subjects who reacted primarily with guilt with subjects who reacted with fear. The guilt-arousing situation would elicit different reactions from the high guilt subjects than from the low guilt subjects. Subjects were then broken down into these two groups depending on their behavior in the guilt-arousing situation. Five subjects who used the "confession" door were placed in the high guilt group. Seven subjects who used the "flight" door were placed in the low guilt group. A subject who showed freezing and one who showed crying combined with indecision and tension were both placed in the high guilt group. A careful study of the behavior protocols of the five remaining subjects was made and each one was individually assigned to the low guilt group or the high guilt group. Consequently, there were a total of 10 children in the high guilt group and a total of 9 in the low guilt group, and when the three children who refused all the stories on the pre-test were excluded this left 8 children in each group.

The matched pair difference scores of the high guilt and the low guilt subjects were then compared by means of the *t* test for independent samples (12) and the Kruskal-Wallis One-way Analysis of Variance by Ranks (18). Only the Change Score means met the assumptions of the *t* test for independent samples. The only difference found approaching significance was in the Change score. The high guilt group had slightly higher change scores relative to their controls than the low guilt group. On dissimilar stories,

this difference was significant at the .11 level. On total stories, this difference was significant at the .10 level. Therefore, most of the difference found between the experimental and the control groups in Change score is probably due to the high guilt group. Consequently, the first two hypotheses were not supported by the results.

The third hypothesis predicted a positive relationship between the child's responses to the deviation doll play stories and his responses to the guilt-arousing situation; that is, the fantasy behavior would be positively related to reality behavior. To test this hypothesis, 1008 four-fold tetrachoric correlation coefficients were computed and tested for significance by Fisher's Exact Test (18). The total number of subjects in each case was 19. All possible scores and combinations of scores derived from the stories were correlated with all scores and combinations of scores based on the behavior in the guilt-arousing situation. Of these 1008 correlations, only seven were significant at the .05 level of significance (two-tail test). In summary, the correlations computed between the story completion responses and behavior in the real life deviation situation failed to demonstrate any relationships between these two sets of responses.

The above results raises the question as to whether the children's behavior is sufficiently consistent to yield any significant relationships. How reliable are the children's responses to the stories? Since the pre-test and the post-test were matched sets of stories the consistency of the child's responses could be determined by correlating his responses to the pre-test with his responses to the post-test. A reliability estimate of this type would be most meaningful for the control group of subjects. When these correlations were computed, the following turned out to be significant: Fixing .66* on similar stories (* .05 level of significance) and .89** on total stories (** .01 level of significance); Physical Punishment, .85* on similar stories, and .72* on dissimilar stories; Other Punishment .94* on dissimilar stories; Total Punishment 1.00** on similar stories, .88* on dissimilar stories, and .99** on total stories.

Since the guilt-arousing experience had no specific effects on the individual scoring categories, the reliability coefficients were also computed for the experimental group. The following scores showed significant reliability: Fixing .90** on similar stories and .80* on total stories; Other Punishment .78* on similar stories and .80* on total stories. The experimental group contained similar types of reliability to the control group, so both groups were combined and the reliabilities were computed. Significance was then

found for the following reliability coefficients: Fixing .87** on similar stories and .74** on total stories; Physical Punishment .70** on similar stories, .80** on dissimilar stories, and .70** on total stories; Verbal Punishment .64** on dissimilar stories and .75** on total stories; Total Punishment .76** on similar stories, .80** on dissimilar stories, and .92** on total stories; Confession .65* on similar stories; Redefinition Plus .75* on similar stories.

It is evident from the above that a large proportion of the story completion categories are not of satisfactory reliability. This is undoubtedly due, in part, to the low frequency of occurrence of many of these responses. Most of the children's responses to the stories involved either Fixing or some form of Punishment; for the total group, 492 out of the total of 678 responses fell into one of these two categories. These are the response categories which also showed a high degree of reliability. Unfortunately, there was no way to determine the reliability of the child's behavior in the guilt-arousing situation.

D. DISCUSSION

The fact that none of the hypotheses were confirmed by the results raises some serious questions concerning either the hypotheses themselves or the validity of the measuring techniques. The insufficient reliability of many of the story response categories would preclude the possibility of obtaining significant relationships, but even those categories that were reliable failed to yield significant findings. The question as to the success of the guilt-to yield significant findings. The question as to the success of the guilt-arousing situation in arousing guilt must be considered. The children's behavior in the situation strongly indicates that they were upset by it: such reactions as crying, flight, confession, and freezing indicate that the guilt-arousing situation did influence the children. The upsetting effects of the guilt-arousing situation were also reflected to some degree in the Change score and in changes in the Refusal category.

The children in the experimental group changed their responses significantly more than did the children in the control group. There is suggestive evidence that this difference may have been produced by the high guilt children. In other words, those children who showed high guilt in their reactions to the real life deviation situation tended to be more apt to alter their responses to deviation stories following the guilt-arousing experience. That the guilt-arousing experience had a greater effect on the high guilt child than on the low guilt child is evidence that it did, in fact, arouse guilt in a certain number of children.

The interpretation that offers the most plausible explanation of the results deals with the question of the adequacy of the thematic stimulus material. The incomplete stories tended to elicit only a few very stereotyped reactions, i.e., fixing and punishment. The stories failed to evoke the wide range of fantasy reactions which undoubtedly exist in a group of children such as were studied. Not enough variability in responsiveness was elicited to make differentiations among the children or to obtain relationships between measures. This limited range is probably due to the amount of structure in the stories. The incomplete story technique as used in this study was too highly structured to allow the subject to make a free response. The stories present too many specific, clear-cut stimuli to the subject. The subject perceives the stories as questions about how he would behave in the situations rather than as free fantasy material.

What type of responses do the stories elicit? The most frequent response is either some form of punishment or fixing, both of which are socially acceptable responses. Within this restricted range of responses the high reliability probably results from the child's using a response which he feels is socially acceptable over and over again. Consequently, the type of incomplete story that was used in this study is unlikely to be effective as a measure of guilt and identification. The results suggest that if doll play incomplete stories are to be used in the study of guilt, or anything else for that matter, the total stimulus situation must be presented to the child in a somewhat less structured form. There is a definite need to study the effect of varying the amount of structure in the stories on the child's completions of the stories.

E. SUMMARY AND CONCLUSIONS

Thirty-eight nursery school children were divided equally into an experimental group and a control group and given a standard 20-minute doll play session to establish rapport. Following this, they were administered a pre-test and a post-test consisting of six matched pairs of incomplete doll play stories all of which involved deviations by the child doll similar in sex to the subject. During the interval between the pre-test and the post-test, the experimental group underwent a guilt-arousing experience and the control group did not. The guilt-arousing situation was the collapsing of a doll house with which the child was playing. The child's completions of the stories and his responses to the collapsing doll house were categorized and recorded.

Hypotheses were (a) that the guilt-arousing situation would produce an

increase in the guilt responses of the experimental subjects to the doll play stories, (b) that this increase would be greatest in the stories similar to the guilt-arousing situation and least in the stories different from it, and (c) that there would be positive relationships between their guilt responses in the doll play situation and their responses to the guilt-arousing situation. The results failed to confirm these hypotheses.

The only differences found approaching significance were as follows: the experimental group showed a greater increase in the refusal of stories and more response change following the guilt-arousing situation than did the control group. Further analysis revealed that the difference in total response change was primarily due to high guilt children as opposed to low guilt children. These results showed that the guilt-arousing situation upset the children, and that the high guilt children were more upset and disorganized by it than the low guilt children.

The most likely explanation of the results was offered in terms of the characteristics of the thematic stimulus material. The doll play stories used in this study were too highly structured so that they elicited only a few, every stereotyped response classes. High reliability was limited to those responses dealing with punishment and fixing. The doll play deviation story does not allow the freedom of expression needed for fantasy responses. The effect of varying the amount of structure in the stories on the child's completions of the stories needs to be studied.

REFERENCES

1. AMMONS, C. H., & AMMONS, R. B. Aggression in doll play: Interviews of two- to six-year-old white males. *J. Genet. Psychol.*, 1953, **82**, 205-213.
2. AMMONS, R. B., & AMMONS, H. S. Parent preferences in young children's doll-play interviews. *J. Abn. & Soc. Psychol.*, 1949, **44**, 409-505.
3. BARUCH, D. W. Aggression during doll play in a preschool. *Amer. J. Orthopsychiat.*, 1941, **11**, 252-260.
4. CHITTENDEN, G. E. An experimental study in measuring and modifying assertive behavior in young children. *Monogr. Soc. Res. Child Devel.*, 1942, **7**, No. 1.
5. HARTLEY, E. L., & SCHWARTZ, S. A pictorial-doll play approach for the study of children's intergroup attitudes. *Int. J. Opin. & Attitude Res.*, 1951, **5**, 261-270.
6. HEINICKE, C. Some antecedents and correlates of guilt and fear in young boys. Unpublished Ph.D. thesis, Harvard University, 1953.
7. HENRY, J., & HENRY, Z. Doll play of Pilaga Indian children; an experimental and field analysis of the behavior of the Pilaga Indian children. *Res. Monogr. Amer. Orthopsychiat. Ass.*, 1944, No. 4.
8. KORNER, A. F. Some Aspects of Hostility in Young Children. New York: Grune & Stratton, 1949.
9. KRALL, V. Personality characteristics of accident repeating children. *J. Abn. & Soc. Psychol.*, 1953, **48**, 99-108.

10. LEVY, D. M. Use of play technique as experimental procedure. *Amer. J. Orthopsychiat.*, 1933, **3**, 266-277.
11. ———, Hostility patterns in sibling rivalry experiments. *Amer. J. Orthopsychiat.*, 1936, **6**, 183-257.
12. MCNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949.
13. MURPHY, L. B. Patterns of spontaneity and constraint in the use of projective materials by pre-school children. *Trans. N. Y. Acad. Sci.*, 1942, **4**, 124-128.
14. ROSENZWEIG, S., & SHAKOW, D. Play techniques in schizophrenia and other psychoses: An experimental study in schizophrenic constructions with play materials. *Amer. J. Orthopsychiat.*, 1937, **7**, 32-47.
15. SEARS, P. S. Doll play aggression in normal young children: Influence of sex, age, sibling status, father's absence. *Psychol. Monog.*, 1951, **65**.
16. SEARS, R. R., MACCOBY, E., & LEVIN, H. *Patterns of Child Rearing*. New York: Row, Peterson, 1957.
17. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, P. S. Some child rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monog.*, 1953, **47**, 135-234.
18. SEIGEL, S. *Non-Parametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
19. STOLZ, L. M. *Father Relations of War-Born Children*. Stanford: Stanford Univ. Press, 1954.

Department of Psychology

Michigan State University

East Lansing, Michigan

AN ASSESSMENT OF THE FAKABILITY OF SCORES ON THE GROUP PERSONALITY PROJECTIVE TEST*

Department of Psychology, Phoenix Union College

RUSSELL N. CASSEL AND ROBERT P. BRAUCHLE

A. PROBLEM

This study is concerned with determining the degree to which a select group of North East High School seniors, San Antonio, Texas are able to fake both good and poor personality patterns on the basis of scores from The Group Personality Projective Test (*GPPT*), and in terms of comparisons with scores from a regular and routine administration of the test.

B. DESCRIPTION OF TEST

The *GPPT* has recently been published and is now distributed by Psychological Test Specialists, Missoula, Montana. It is purported by the authors to be relatively free of fakability by individuals desiring to disguise their scores in a direction that might prove of vantage to them. Each of the 90 items on the test includes a "stick-figure" type drawing in which there is a paucity of detail present, and five multiple choice type answers based on factorial validity.

Theory underlying the test presumes that active need areas present in the subject's middle layers of the personality (the mask the individual wears in relation to self and select intimate friends) will be utilized in completing the meaning to the referenced drawings, and which is desirable for selecting one of the five multiple choice answers. Each of the subject's choices of an answer as to what is taking place in the partially unstructured drawing is presumed to be a self projection of the subject of active and unsatisfied need or tension areas in the personality for the time of the test. Through a compounding of the manifest need areas a total score is derived which is presumed to be indicative of total personality disturbance.

The six part scores on the test are described by the authors as follows: (a) tension reduction quotient—percentage of negative feelings in relation to the total number of positive and negative feelings projected; (b) nurturance needs score—need to play the father rôle and give aid; (c) withdrawal needs

* Received in the Editorial Office on June 23, 1958.

score—need to escape; (*d*) neuroticism needs score—inability to make decisions; (*e*) affiliation-psychosexual needs score—desire to belong and for boy-girl relations; and (*f*) succorance needs score—need to play mother rôle and seek aid.

C. DESCRIPTION OF GROUP TAKING TEST

Fifty high school seniors, about equally divided between boys and girls, and all members of home room groups of the high school counselor, and all of whom volunteered to participate in the research experiment (rather enthusiastically). They ranged in age from 15 to 19 years old with a *M* age of 17.1, and a *SD* of 0.71 years. In terms of *IQ* scores from group administered regular school tests their intelligence ranged from 80 to 126 with a *M* of 106.7, and a *SD* of 18.43 *IQ* points.

The three administrations of the *GPPT* were made by the head counselor during the early part of the 1957-58 school year. The initial and regular administration of the test was accomplished as a portion of the regular school test program, and no additional explanation was made at the time of the test.

D. EXPERIMENTAL ADMINISTRATIONS OF TEST

Two home room groups of the senior class, both of which were assigned to the referenced counselor, were invited to participate in a psychological experiment designed to determine the degree to which high school subjects could fake good and poor personality patterns as indicated by scores on the *GPPT*. All of the members of both home room groups invited volunteered rather enthusiastically to participate in the experiment. There was no reference made by the counselor as to what constituted either a good or poor personality pattern, but there was considerable discussion among the students on the point.

The second administration of the test was structured to fake a poor and highly disturbed personality pattern. The students were asked to study the five choices provided for each of the 90 items and to select the one they believed would make their scores reveal a poor and highly disturbed personality pattern. This administration was made several days after the initial and regular administration of the test was made.

The third administration of the test was structured to fake a good personality pattern with a minimum of disturbance present, and it was given about two weeks after the initial administration of the test.

E. FINDINGS

The findings from this study are illustrated in Table 1, and they are all based on comparisons of the part and total scores for the initial or regular administration of the *GPPT*, and the two subsequent experimental administrations (designed to fake a poor and a good personality pattern).

Three types of scores appear to be meaningful and significant in making the comparisons between the initial or regular administration and the two experimental ones: (a) scores that favor the intended faked direction, (b) scores that oppose the intended faked direction, and (c) scores that are not significantly different in the two administrations, i.e., regular and experimental.

1. *Faked Poor Personality Scores*

In general, the data indicates that the *GPPT* can be faked very effectively in a direction indicating a poor personality pattern.

a. *Scores that favor intended faked direction.* Five of the six part-scores and the total score all show significant vantage in terms of the indicated direction consistent with the intended faking (poor personality pattern): (a) higher tension reduction quotient, (b) lower nurturance needs score, (c) higher withdrawal needs score, (d) lower affiliation-psychosexual needs score, (e) higher succorance needs score, and (f) higher total score (greater personality disturbance).

b. *Score showing no difference in two administrations.* Only the neuroticism needs score showed no difference between the regular administration and the administration designed to fake a poor and highly disturbed personality pattern. For this administration, none of the scores opposed the intended direction of the faking.

2. *Faked Good Personality Scores*

In general, the data indicate that the 50 students were unable to fake scores on the *GPPT* in a direction indicating a good personality pattern, and with a minimum of disturbance present.

a. *Scores that favor intended faked direction.* Three of the part scores showed significant agreement in terms of vantage for the subject for the intended faking of the test, i.e., good personality pattern: (a) lower tension reduction quotient, (b) higher nurturance needs score, and (c) lower succorance needs score.

b. *Scores that oppose intended faked direction.* Three part-scores on the

TABLE 1
An Illustration of the Reliability of Difference between *M*'s and *SD*'s for Three Separate Administrations
(Regular, Faked, Poor Personality, and Faked Good Personality) of the *GPPT* to a Select Group of 50
North East High Seniors, San Antonio, Texas

Scores & conditions	Tension reduction quotient	Nurtur- ance needs score	With- drawal needs score	Neuroti- cism needs score	Affiliation psychosexual needs score	Succor- ance needs score	Total personality needs score
(1) Regular administration mean	29.24	35.30	119.70	45.67	21.90	87.40	59.26
Standard deviation	7.44	9.57	31.80	11.23	6.63	58.00	9.24
(2) Faked poor personality mean	49.56	15.30	252.30	45.90	7.87	263.66	84.00
Standard deviation	7.92	12.57	160.00	11.93	4.00	59.30	11.42
(3) Faked good personality mean	25.32	45.53	136.00	70.90	18.20	38.13	56.04
Standard deviation	6.78	11.17	29.40	16.97	5.27	27.03	10.10
Diff. in <i>M</i> 's for (1) & (2)	20.32	-20.00	132.60	0.23	-14.03	176.20	24.74
<i>t</i> Value for dif. (1) & (2)	**11.75	**7.14	**4.16	0.07	**8.16	**5.40	**10.76
Diff. in <i>M</i> 's for (1) & (3)	-4.32	8.23	16.30	25.23	-3.70	-49.27	-3.22
<i>t</i> Value for dif. (1) & (3)	**2.82	**3.08	*2.09	**7.29	*2.30	**4.69	1.61
Diff. in <i>SD</i> 's for (1) & (2)	0.48	3.00	128.20	0.70	-2.63	1.30	2.18
<i>t</i> Value for dif. (1) & (2)	0.39	1.44	**5.60	0.29	*2.14	0.90	1.35
Diff. in <i>SD</i> 's for (1) & (3)	-0.66	1.60	-2.40	5.74	-1.36	-30.07	0.86
<i>t</i> Value for dif. (1) & (3)	1.11	0.82	0.43	*2.31	1.17	0.52	0.54

* Statistically significant at 5 per cent level.

** Statistically significant at 1 per cent level.

Note: All part scores are multiplied times their respective "B" weights, and as described in manual.

test show significant disadvantage in terms of disagreement with the desired direction of faking, i.e., good personality pattern: (a) higher withdrawal needs score, (b) higher neuroticism needs score, and (c) lower affiliation-psychosexual needs score.

c. *Score showing no difference in two administrations.* There was no significant difference for the total score on the two administrations of the test. Accordingly, the *GPPT* was not fakable in a direction indicating less disturbance and needs of an undesirable type.

F. CONCLUSIONS

For the 50 subjects involved in this study, the *GPPT* was effectively faked in a direction to indicate a poor personality in relation to the regular personality pattern, and in terms of high tension, low desirable needs, and high undesirable needs.

The same individuals were unable to fake the *GPPT* effectively in terms of an overall score indicative of a good personality pattern, or a better pattern in relation to a regular pattern, and in terms of lower tension, higher desirable needs, and lower undesirable needs.

G. SUMMARY

This study is concerned with determining the degree to which 50 select high school seniors are able to fake good and poor personality patterns on the Group Personality Projective Test (*GPPT*) in relation to the pattern established by a previous and regular administration of the test. The findings indicate that the students were able to fake a poorer personality pattern most effectively in all of the part scores but one (neuroticism needs score), and on the total score. They were unable to fake a better total score, however, indicating that the *GPPT* can not be faked to indicate a better or less disturbed personality pattern than is actually present. The projected notions of the subjects attempting to fake a poor and highly disturbed personality pattern were in excellent agreement with the theory underlying the test, i.e., high tension, low nurturance, high withdrawal, high neuroticism, low affiliation-psychosexual needs, high succorance, and high total needs and tension. The projections of the subjects attempting to fake a good personality pattern were partially in agreement with the theory of the test, but were equally in disagreement with that theory, i.e., moderate tension, high nurturance, low withdrawal, low neuroticism, high affiliation-psychosexual needs, low succorance, and low overall tension and needs present.

Since the *GPPT* cannot be faked in a direction to indicate a better personality pattern, it should prove of extreme value for making diagnostic assessments of the middle layer of an individual's personality pattern (the mask worn in relation to self and select intimate friends).

REFERENCES

1. BURKHOLDER, D. Validating a battery of psychological tests for use in the counseling of Latin students. Unpublished master's thesis, Our Lady of the Lake College, San Antonio, Texas, 1956.
2. CASSEL, R., & HARRIMAN, L. A comparative analysis of personality and ego strength test scores for in-prison, neuro-psychiatric, and typical individuals. *J. Exper Educ.* (pending).
3. CASSEL, R., & KAHN, T. C. Development and standardization of the Group Personality Projective Test. *J. Projec. Tech.*, Vol. 22, No. 3, 1958.
4. HARTMAN, B. Developing a functional definition of ego strength by factor analyzing available tests. Unpublished master's thesis, Our Lady of the Lake College, San Antonio, Texas, 1958.
5. JACQUES, O. M. Predicting juvenile delinquency proneness by group tests. *Person. & Guid. J.*, March, 1958.

17187 Hawthorne Avenue
Fontana, California

Department of Counseling
North East High School
San Antonio, Texas

MANIPULATION OF VISUALLY HOMOGENEOUS STIMULUS OBJECTS*

Department of Psychology, University of Texas

RICHARD M. CARR AND W. LYNN BROWN

A. INTRODUCTION AND PROBLEM

Harlow, Blazek, and McClearn (4) have found that manipulatory behavior generally increases in amount and efficiency with age and practice. They suggest that manipulatory behavior is self-sustaining and is not dependent upon, nor derived from internal drives such as hunger or thirst, or their incentive systems.

Recent experimentation by Welker (6) working with chimpanzees has demonstrated that the chimpanzee will preferentially explore and play with those stimulus objects which are, within limits, larger, brighter, and more heterogeneous. Carr and Brown (1, 2) report that monkeys show a preference for the materials of which stimulus objects are constructed; the order of preference being wood, rubber, cork, plastic, and metal. Since the objects used were selected at random for the file of multidimensional objects of the laboratory and differed not only in materials but also in size, color, and shape, it is possible that factors other than those found in material differences could have determined the results.

The present investigation was designed to study the effect of the removal of visual differences (size, color, and shape) among the stimulus objects upon the manipulatory preference choices of the rhesus monkey. The results obtained by Welker (6), Carr and Brown (1, 2), Cho and Davis (3), and McDowell (5) upon manipulatory preference has led the authors to hypothesize that the observed preferences are due to the heterogeneity of the stimulus objects. This experiment was designed to control for visual differences, thus determining the differences in manipulation due solely to tactile heterogeneity.

B. METHOD

1. Subjects

The Ss were 10 experimentally naïve rhesus (*Macaca mulatta*) monkeys approximately four years of age. The Ss were fed at 4:00 P.M. daily.

* Received in the Editorial Office on June 26, 1958.

2. *Apparatus and Procedure*

The *Ss* were placed in a metal cage measuring 2 x 3 x 3 ft. A special manipulation box containing three stimulus objects constructed of the same material was mounted on the barred front of the cage. The manipulation box was constructed of 3/8 inch plywood and measured 28 x 12 x 4 inches. The front of the box contained three 1/4 inch holes, one in the center and the others 7 inches from each side. Springs were inserted in these holes and attached to an automatic recording device within the box at one end and to the stimulus objects at the other end. A pressure of 50 grams upon any of the stimulus objects would activate the recording device which recorded the number of manipulations as well as the amount of time spent manipulating the stimulus objects.

The *Ss* were tested for five 5-day periods; the duration of each daily session was 10 minutes. A latin square design with two replications was employed so that for each group of 5-day periods two animals were presented with the same three cylindrical stimulus objects one inch in length and one inch in diameter, painted red in color, and weighing 22 grams. Five sets of three objects were employed; the sets were constructed of wood, metal, rubber, cork, and plastic, respectively. In this manner all visual differences between objects were minimized and preference for any particular material would be made solely upon the basis of tactile differences of malleability and possibly the temperature of the stimulus objects. Each group of *Ss* was presented with three stimulus objects constructed of the same material for a five-day period. Records of the number of manipulations and the amount of time spent manipulating at each session were kept for each *S*. All *Ss* were run in close succession each day with the order of testing the *Ss* being randomized.

C. RESULTS

An analysis of variance applicable to a latin square design was employed as an over-all test of significance. The analysis demonstrated that there was no significant preference in choice of nor in time spent manipulating stimulus objects. However, the amount of manipulation and time spent manipulating the stimulus objects over the five-day testing sessions was significant at the .01 and .025 levels respectively. This demonstrates that manipulatory behavior steadily increased through the first four of the five-day sessions and then dropped sharply in the fifth week (Figures 1 and 2).

D. DISCUSSION

The results of this experiment demonstrate that when monkeys are presented with sets of visually homogeneous stimulus objects: (a) the Ss no longer show a preference for any particular set of such objects, and (b) manipulation continues to increase for a period of four weeks and then dropped sharply in the fifth week.

Since the sets of stimulus objects differ only in terms of tactile stimula-

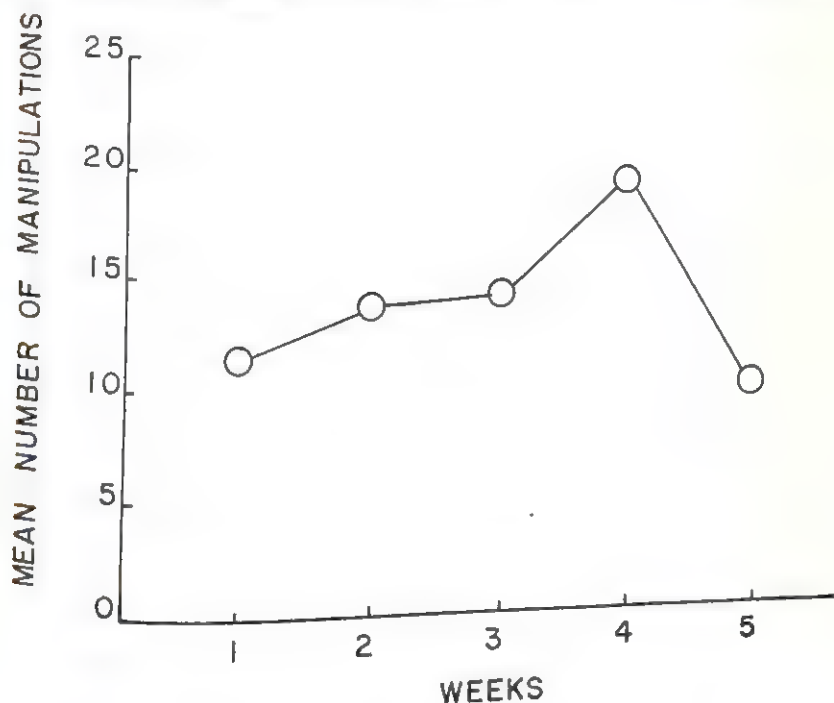


FIGURE 1
MEAN NUMBER OF MANIPULATIONS OF THE STIMULUS OBJECTS BY ALL Ss FOR EACH OF THE FIVE WEEKS OF PRESENTATION

tion, i.e., malleability and possibly temperature, the monkey no longer shows a preference for one set above the others as is found in the work of Welker (6) and Carr and Brown (1, 2). The preferences found in previous experiments (1, 2, 3, 5, 6) are therefore hypothesized as being due to one or more differences in the visual heterogeneity of the stimulus objects, i.e., color, size, and shape.

The experimental results support the hypothesis of Harlow, Blazek, and McClearn (4) and Carr and Brown (1, 2) that manipulatory behavior is

self-sustaining and is independent of internal drives such as hunger and thirst (4) as well as the hypothesis that the novelty of the situation is self-rewarding (1). However, when the monkey is presented with homogeneous visual stimuli, the novelty of the situation decreases more rapidly than when the *S* is presented with visually heterogeneous objects and consequently the incentive motivation produced by novelty decreases more rapidly.

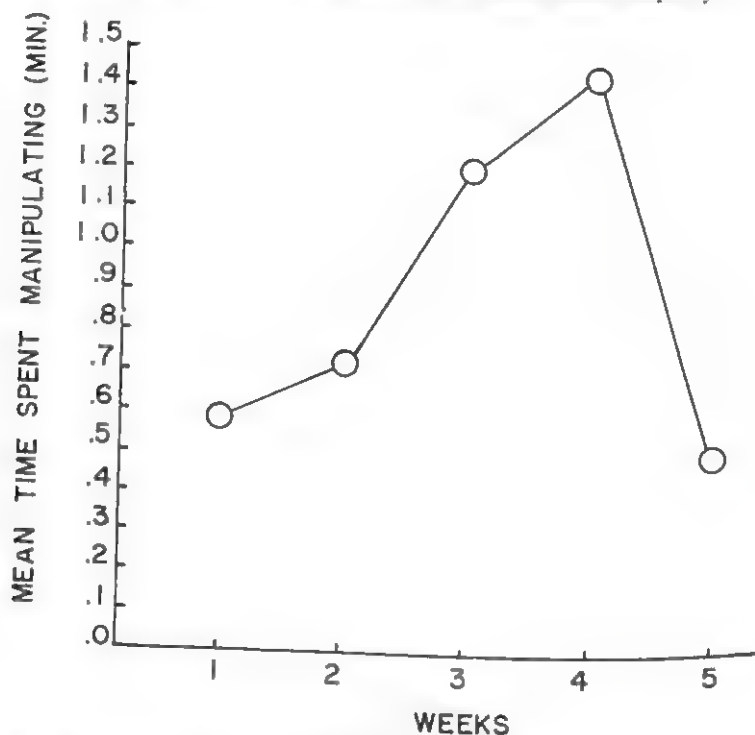


FIGURE 2
MEAN TIME SPENT MANIPULATING THE STIMULUS OBJECTS BY ALL *Ss* FOR EACH OF THE FIVE WEEKS OF PRESENTATION

E. SUMMARY

Ten experimentally naïve male monkeys were tested on a special manipulation box for 10 minutes each day for five days of each of five different stimulus objects of the same size, shape, color, and weight. The objects were made of wood, metal, cork, rubber, and plastic.

The results demonstrate that monkeys, when presented with sets of visually homogeneous stimulus objects, no longer show a preference for any particular set of objects and that manipulation will continue to increase over a period of four weeks and will then rapidly decline.

REFERENCES

1. CARR, R. M., & BROWN, W. L. The effect of the introduction of novel stimuli upon manipulation in rhesus monkeys. *J. Genet. Psychol.*, 1959, **94**, 107-112.
2. ———. The effect of sustained novelty upon manipulation in rhesus monkeys. *J. Gen. Psychol.*, 1959, **61**, 121-125.
3. CHO, J. B., & DAVIS, R. T. Preferences of monkeys for objects other than food. *Amer. J. Psychol.*, 1957, **70**, 87-91.
4. HARLOW, H. F., BLAZEK, N. C., & MCCLEARN, G. E. Manipulatory motivation in infant rhesus monkeys. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 444-448.
5. McDOWELL, A. A. Normative behavior of the cage Macaque (*Macaca mulatta*). Unpublished Master's Thesis, University of South Dakota, 1953.
6. WELKER, W. L. Some determinants of play and exploration in chimpanzees. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 84-89.

Department of Psychology
University of Texas
Austin, Texas

OBSERVATIONAL LEARNING OF AN AVOIDANCE RESPONSE*¹

Department of Psychology, Emory University

W. J. PRESLEY AND A. J. RIOPELLE

A. INTRODUCTION

Numerous attempts have been made to determine whether the learning of a response by one animal is enhanced after observing the performance of another. Adler (1) has recently reviewed the literature and has suggested that, although results with the same species are often contradictory, some evidence for the phenomenon has appeared in most of the studies. In light of his review, no further discussion on the general subject is needed.

Most of the investigations on the problem have used food reward. Brogden (2), however, used shock motivation in a study on imitation and social facilitation in the conditioning of forelimb-flexion in dogs. Twenty mongrel dogs were assembled into pairs. Both animals in each pair were placed in conditioning stocks. One animal was conditioned while the other had an opportunity to observe. Later the observer was conditioned. No evidence for enhancement of learning by the observation was found.

Perhaps the crucial response was not obtrusive enough, or perhaps the observers were not attentive to the demonstrators. Brogden reports that some observers became lethargic and others struggled to get out of the stock.

In the present experiment we tried to investigate the same phenomenon but in a different fashion. We wanted the demonstrator to make a more obtrusive response. In addition, we avoided restraining the observer. Finally, we used a primate, the rhesus macaque. We chose this species because it pays close attention to other monkeys.

B. SUBJECTS

Ten adolescent rhesus monkeys participated in the experiment. None of the animals had served in any prior experiment.

* Received in the Editorial Office on July 2, 1958.

¹ This research is supported in part by research grant (No. M-589) from the National Institute of Mental Health of the National Institutes of Health, U.S. Public Health Service.

C. APPARATUS

A double-compartment cage was used. The two compartments were 3 ft. x 3 ft. x 3 ft. in dimensions and were separated by a transparent plastic screen. Plastic screening served as a wall on the experimenters side. It was covered by a cloth to provide one-way vision.

Floors of the compartment were gratings built of 3/8-in. iron rods spaced 2 in. on center.

Erected through the center of the cage was a barrier 12 in. high. The grating on each side of the barrier could be electrified independently with 5 ma. of pulsed current. The floor of *O*'s compartment was never electrified.

The interior of the cage was illuminated from above by a white, 75-watt lamp. Beneath the grating and on each side of the barrier were two light bulbs, each 75-watts, painted red. The lights in *D*'s compartment were turned on and off by the experimenter, those under *O* were never turned on.

D. PROCEDURE

One monkey was placed in each compartment. At the beginning of every trial the red light beneath the monkey being trained (Demonstrator) was turned on. It remained on for 14 sec. Four seconds after the light was turned on, the grid in the lighted compartment was electrified for 10 seconds. Shock from this grid could be avoided by the monkey by leaping over the barrier within the first four seconds after the onset of the light. On the next trial, 10-30 sec. later, the demonstrator was shuttled back to the original side. The observer in the adjacent non-electrified compartment could observe the entire procedure.

E. EXPERIMENTAL DESIGN

The basic unit of the experiment was a pair of animals. One member of the pair served as Demonstrator (*D*) and the other as Observer (*O*). While *D* was learning the reaction, *O* observed from the adjacent compartment. Sessions of 60 trials each were given until *D* made 28 avoidance responses in a block of 30 trials. When *D* reached this criterion, his performance in the experiment ended. *O* was then transferred (in the next session) to the electrified side of the apparatus. Testing continued as with *D*.

To check on the possibility that the greater familiarity with the apparatus on the part of *O* might lead artifactiously to better performance, three *D*s were placed into the apparatus for three periods of one hour each prior to training. The remaining two *D*s received no familiarization training.

F. RESULTS

Because the familiarization sessions did not materially alter original learning ($t = .13$), performances for the two subgroups of *D* were combined. The comparison of the learning by *Ds* and by *Os* is shown in Figure 1. The

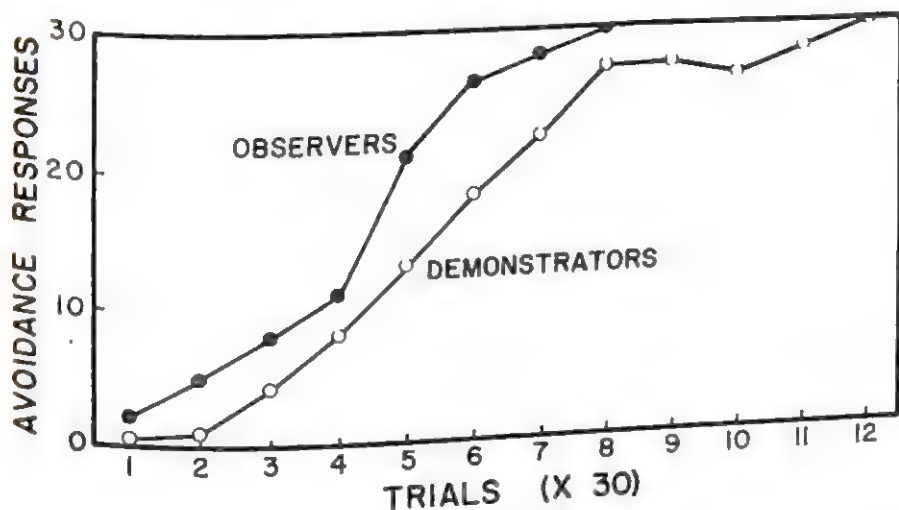


FIGURE 1
THE LEARNING OF AN AVOIDANCE RESPONSE BY OBSERVERS AND BY DEMONSTRATORS

Os' performance was superior to that of the *Ds* at all stages of training. The over-all difference between groups was significant at the 5 per cent level ($t = 2.9$). The slowest *O* learned the problem in as few trials as did the fastest *D*. These data thus show that rhesus monkeys learn an avoidance response more rapidly after having observed another learn it than they would have without this observational experience.

G. DISCUSSION

These results contrast with previous findings in two particulars. First, previous meaningful demonstrations of observation learning have used rewards as incentives. The present data thus extend previous data. Secondly, the one prior attempt, by Brogden, did not obtain observational learning in dogs. Whether the difference in results is due to generic or methodological differences is not known with certainty. We do know, however, that with appropriate procedure, positive effects can be obtained with cats (1, 3).

H. SUMMARY

Rhesus monkeys were trained on an avoidance response after having watched another monkey learn the response. When trained later, the observers learned faster than did the demonstrators.

REFERENCES

1. ADLER, H. E. Some factors of observational learning in cats. *J. Genet. Psychol.*, 1955, **86**, 159-177.
2. BROGDEN, W. J. Imitation and social facilitation in the social conditioning of forelimb flexion in dogs. *Amer. J. Psychol.*, 1942, **55**, 77-83.
3. HERBERT, M. J., & HARSH, C. M. Observational learning by cats. *J. Comp. Psychol.*, 1944, **37**, 81-95.

Yerkes Laboratories
Orange Park, Florida

EFFECT OF EXOGENOUS ADRENALIN ON THE STARTLE RESPONSE IN THE RAT*

Department of Psychology, Carnegie Institute of Technology

K. E. MOYER AND B. N. BUNNELL¹

A. INTRODUCTION AND PROBLEM

Brown, Kalish, and Farber (3) have shown that the magnitude of the startle response in rats is increased when the startle stimulus is presented in conjunction with a light and buzzer which previously had been followed by shock. They suggested that the previous pairing of light and buzzer with shock caused the animals to develop a conditioned fear reaction. It has been suggested that the secretions of the adrenal gland, particularly adrenalin, play a significant rôle in the production and maintenance of the physiological components of fear (4, 7). Adrenalin is regarded as a sympatheticomimetic. It is said to be liberated by the adrenal medulla during strong emotion and thereby perpetuates the sympathetic components of the emotion. This point of view has been severely criticized by Rogoff (11) and recent animal experiments have failed to demonstrate the expected relationship between adrenal functioning and fear motivated behavior. Applezweig and Moeller (2) have failed to find any differences between normal and bilaterally adrenalectomized rats in the rate of learning an avoidance response. Moyer (8) has also demonstrated that adrenalectomy does not adversely affect the learning or retention of an escape response. Moyer and Bunnell (10) have shown that the administration of exogenous adrenalin has no effect on the learning of an avoidance response and that adrenal demedullation does not effect the acquisition of the same response (9). Kosman and Gerard (6), using escape and avoidance learning in a Skinner box, found that large doses of subcutaneous adrenalin cause a decrease in frequency of both a conditioned avoidance response and an escape response.

Applezweig and Moeller (2) specifically studied the effect of adrenalectomy on the capacity of rats to acquire a conditioned fear response as manifested by increased sensitivity to startle stimulation and found that the capacity was not impaired. Although adrenalectomy does not interfere with

* Received in the Editorial Office on July 14, 1958.

¹ This investigation was supported in part by a research grant, M-1298, from the National Institute of Health, Public Health Service.

the acquisition of a conditioned fear response, it is possible that injected adrenalin may duplicate the internal state of the organism during fear. In fact, Landis and Hunt (7) provoked both "Cold" and true emotion by the injection of adrenalin into humans.

If adrenalin is an important physiological component of conditioned fear, and rats in a state of conditioned fear show increased startle reactions, we might expect rats to whom exogenous adrenalin has been administered to exhibit exaggerated startle responses. It is the purpose of this experiment to test this possibility.

B. METHOD

1. *Subjects*

In this study 40 naïve male hooded rats from the colony of the Carnegie Institute of Technology were used. All Ss were between 74 and 77 days of age at the time of testing. Litter mate control was used.

2. *Apparatus*

The apparatus used was a slightly modified form of the one described by Brown, Kalish, and Farber (3). As in their apparatus, the sensing unit was a postage scale (Hanson Model 1509, 5-lb. capacity). The platform, dial, pointer, and protective glass were removed. The platform was replaced with a box $2\frac{3}{4}$ " x 6" x 4" with plastic sides and masonite ends. The top was also of masonite. The floor of the box was composed of brass rods $\frac{1}{16}$ " in diameter spaced $\frac{1}{4}$ " apart. When a S was placed in the confining box and subjected to a sharp auditory stimulus, the startle reaction of the animal caused the platform to be depressed. The downward movements of the platform were transmitted by a thread and pulley system to lever recording pen on a Phipps and Bird moving paper recorder. The recorder was set so that the paper moved at a rate of .15 in. per sec. The startle stimulus was produced by a toy pistol (Super Nu-Matic Paper Buster Gun) held 6" over the confinement box.

3. *Procedure*

Four groups of 10 animals each were used. Group I received .06 mg. of adrenalin chloride per 100 gms of body weight; Group II received .03 mg.; Group III received physiological salt solution; and Group IV received no injection. Parke Davis adrenalin chloride 1:1000 preparation was used. It was diluted with physiological salt solution and a solution was made up for each group so that .1 cc of solution contained the proper amount of adrenalin

for that group, i.e., .1 cc of solution for Group I contained .06 mg. of adrenalin and Group II, .03 mgs. Group III was given .1 cc of physiological salt solution per 100 gm. of body weight. Thus each *S* was injected with the same amount of solution according to body weight. Each *S* was weighed daily and given an intraperitoneal injection of the appropriate solution immediately before the day's trials.

Immediately after the injection the *S* was placed in the confining box of the startle apparatus and was given a period of 15 sec. of acclimatization time. The paper buster gun was then held 6 in. above the cage and fired. This initial shot served to settle the animal down. The *S* characteristically ceased all random movements and remained quite still on all four feet in the cage and did not change position appreciably during the rest of the trial. Five more shots were fired at intervals of 23 sec. and the amount of startle was automatically recorded.

C. RESULTS

The size of the startle response to the first experimental shot was measured and the mean magnitude for the various groups was as follows: .06 adrenalin groups 15.10 mm., .03 adrenalin group 12.40 mm., salt control group 13.95 mm., and control 10.95 mm.

A Bartlett's test (5) for homogeneity of variance was applied to these data. The Chi Square value was 1.50 and was not significant at the .05 level of confidence. An analysis of variance of the same scores indicated that the differences between the means of the four groups were not greater than might be expected by chance ($F = .247$).

In order to test the possible influence of the administration of adrenalin on the extinction of the startle response, the size of the startle response to the fifth shot was subtracted from the size of the startle response to the first shot after the one used to settle the *Ss* down. A Bartlett's test of these different scores was not significant at the .05 level $\chi^2 = 1.25$ indicating that the variances were homogeneous. An analysis of variance of the same data also lacked significance ($F = .50$).

Since there were no differences among groups, the data from all of the groups were combined in order to test whether or not extinction of the startle response actually did occur. A *t* test indicated that the first measured startle response was significantly larger than the last one measured. $t = 4.35$ sig. at .001 level of confidence.

D. DISCUSSION

The results of this experiment indicate that exogenous adrenalin in the dosage used apparently has no influence on the magnitude of the startle response in the rat. Further, it does not delay the extinction of this response. Evidently then, we cannot duplicate or approximate the internal fear or anxiety state in the rat by the injection of adrenalin. These findings appear to lend support to the previous findings of Kosman and Gerard, Moyer, and Moyer and Bunnell.

There are at least two possible interpretations of these results. It may be that endogenous adrenalin is an important part of the conditioned fear response and we are not able to duplicate its action with exogenous adrenalin. Sharrer (1, pp. 167-168) indicated that adrenalin liberated by the medulla is liberated with a carrier which releases the adrenalin molecule slowly and that this situation is difficult if not impossible to duplicate. A second possibility is that neither endogenous nor exogenous adrenalin is an essential feature of the conditioned fear or anxiety response. Several investigations now have shown that adrenalectomy does not interfere with presumably anxiety motivated responses, and that adrenalin does not facilitate them.

All of this work to date, however, has been done on the rat, and it must be recognized that these results may be species specific for this subject. Species specificity in regard to hormone action is well recognized (12).

E. SUMMARY

Forty male hooded litter mate rats were used in this study. They were divided into four groups of 10 Ss each. Group I was given an intraperitoneal injection of .06 mg. of adrenalin chloride per 100 gms. of body weight; Group II was given .03 mg.; Group III was given a comparable amount of physiological salt solution; and Group IV was used as a control and given no injection. Immediately after the injection the S was placed in a stabilimeter type of apparatus which recorded the magnitude of the startle reaction to the shot from a toy pistol.

The results indicate that the magnitude of the startle response is not increased by the intraperitoneal injection of adrenalin and that adrenalin has no influence on the extinction of the startle response. The results also show that there is a significant reduction in the magnitude of the startle response over a series of five trials spaced 23 sec. apart.

REFERENCES

1. ABRAMSON, H. A. *Neuropharmacology*. New York: Josiah Macy, Jr. Foundation, 1956.
2. APPELZWEIG, M. H., & MOELLER, G. Hormonal influences in learning: The pituitary-adrenal system, anxiety, and avoidance learning. Office of Naval Research: Research report on Project NR 154-103 and Project NR 154-137.
3. BROWN, J. S., KALISH, H. I., & FARBER, I. E. Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *J. Exp. Psychol.*, 1951, **41**, 317-328.
4. CANTRIL, H., & HUNT, W. A. Emotional effects produced by the injection of adrenalin. *Amer. J. Psychol.*, 1932, **44**, 300-307.
5. EDWARDS, A. L. *Experimental Design in Psychological Research*. New York: Rinehart, 1950.
6. KOSMAN, M. E., & GERARD, R. W. The effect of adrenalin on a conditioned avoidance response. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 506-508.
7. LANDIS, C., & HUNT, W. A. Adrenalin and emotion. *Psychol. Rev.*, 1932, **39**, 467-485.
8. MOYER, K. E. Effect of adrenalectomy on anxiety motivated behavior. *J. Genet. Psychol.*, 1958, **92**, 11-16.
9. MOYER, K. E., & BUNNELL, B. N. Effect of adrenal demedullation on an avoidance response in the rat. (In preparation.)
10. ———. Effect of injected adrenalin on an avoidance response in the rat. *J. Genet. Psychol.*, in press.
11. ROGOFF, J. M. A critique on the theory of emergency function of the adrenal glands: Implications for psychology. *J. Genet. Psychol.*, 1945, **32**, 249-268.
12. SELYE, H., & HEUSER, G. Fifth annual report on stress, 1955-56. Montreal, Canada: Acta, Inc., 1956.

Department of Psychology
Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania

THE EFFECTS OF EARLY ENFORCED WEANING ON SUCKING BEHAVIOR OF PUPPIES*

Roscoe B. Jackson Memorial Laboratory, Bar Harbor, Maine

J. P. SCOTT, SHERMAN ROSS, AND ALAN E. FISHER¹

A. INTRODUCTION

Sucking behavior has attracted considerable scientific interest because of its obvious relationship to the phenomenon of early socialization in mammals, because of the supposed effects of different weaning practices upon the development of personality in different human cultures, and because thumb-sucking in a mildly neurotic form is a frequent occurrence in children of our own society.

The scientific literature on sucking in both human and infrahuman subjects has been reviewed in a recent paper (8). The present research is concerned chiefly with the effects of sudden weaning on puppies of different ages. Age at weaning of human infants is an important factor in determining the amount of emotional disturbance, the greatest effect occurring between 13 and 18 months of age (11). In puppies, which have a shorter developmental time and in which weaning normally occurs from seven to 10 weeks of age, there is a possibility of defining the time element more definitely and relating it to the critical period for socialization (9).

The results with puppies indicate that there is a change in the reaction to weaning correlated with the onset of the critical period for socialization. Another important factor affecting the reaction of puppies to weaning is the previous state of nutrition as indicated by body weight. Finally, comparison with other experimental studies indicates that the nature of post-weaning care is another important factor.

B. SUBJECTS AND MATERIALS

A total of 14 litters and 64 animals were used in the experiment. Eleven

* Received in the Editorial Office on July 17, 1958.

¹ The experiments described in this paper were supported by the United States Public Health Service, National Institute of Mental Health, through Grants M-522 and M-123. The work was done under the joint direction of the two senior authors by Alan E. Fisher, who collected the data for Litters VI-XIV and made the preliminary data analysis, and David J. King, who obtained data on the first five litters. The authors wish to thank John A. King and Orville Elliot for permission to include additional data from their experiment.

litters were cocker spaniels and three were African basenjis. There were four litters of three, four litters of four, and six litters of six. The puppies were weaned at ages ranging from 10 to 36 days after birth, as shown in Table 1.

The general method was to wean the puppies, two at a time, so that isolation would not be involved and so that they would have an opportunity for body sucking if this should arise. The remaining animals were left as controls, except in the litters of six, where two pairs of puppies were usually removed at different ages. After two weeks, the weaned puppies were returned to their mothers. The larger animals in the litter were chosen for weaning, partly because they could better stand the stress of weaning and partly to weight the experiment against the hypothesis that nutritional state would produce sucking. The effect was, of course, to leave an additional milk supply for the control puppies remaining with their mothers. As Figure 4 shows, the result was an upturn in their weight curves.

In three litters, (Nos. 7, 10, 13), all animals were less than 90 per cent of normal weight prior to weaning, indicating that the nutrition supplied by the mother was unsatisfactory. Litter 13 was in particularly poor condition, and all members were weaned in order to increase the chances of survival, but three out of four nevertheless died. There was no other mortality in the course of the experiment.

All experiments were done in the same experimental room. This was approximately 10 x 18 feet in size with a large window, insulated walls, and an exhaust fan for ventilation. Observations were made through a one-way vision mirror. Few disturbing noises could be heard inside the room. Heat was provided by controlled hot air. However, the control mechanism was not efficient, and the temperature varied from 67° to 87° F during the course of the experiment.

The room itself was divided into two identical halves, each with a nest box, food stand, and food and water dishes. The concrete floor was kept covered with shavings, and feces were removed once per day. In most cases, a pregnant female was brought into one-half of the room and had her puppies there. At the time of weaning two puppies were removed to the opposite half of the room. The partition was made of a frame covered with welded wire on both sides, so that the puppies could not come into contact with animals on the other side but could easily see them.

Before weaning, the mother and pups were given the usual routine care and feeding. Dry Purina kibbled dog meal and laboratory chow were avail-

TABLE 1
AGE AT WEANING, RELATIVE WEIGHT, AND SUCKING REACTIONS OF EXPERIMENTAL AND CONTROL PUPPIES

Age at weaning	Litter No.	Litter size at weaning	Experimentals				Controls		
			Wt. before weaning % normal	Times tested during Sep.	Avg. sucking days 2-5	Avg. sucking days 1-14	Wt. before weaning % normal	Avg. sucking days 2-5	Avg. sucking days 1-14
10	14	6	99	10	4.00	4.40	66	1.00	.42
10	14	6	96	10	5.00	5.10	79	1.00	.42
11	13	4	58	13	5.75	5.62	—	—	—
11	13†	4	60	13	4.50	5.00	—	—	—
12	11*	6	115	10	1.25	1.00	88	.00	.00
12	11*	6	115	10	1.00	.70	112	.00	.00
14	9	3	117	12	2.50	.83	98	.00	.42
14	9	3	114	12	.50	.16	—	—	—
14	13†	4	46	13	5.00	5.00	—	—	—
14	13†	4	73	5	6.00	4.80	—	—	—
15	10	3	54	7	4.33	4.44	63	.25	.18
15	10	3	58	8	4.75	3.75	—	—	—
15	14	6	83	11	1.66	.90	65r	.00	.00
15	14	6	78	11	3.33	2.24	75r	.00	.00
16	8	3	84	11	5.25	3.36	62	.75	.55
16	8	3	91	11	4.75	3.27	—	—	—
18	2	4	97	11	.66	1.63	98	.00	.00
18	2	4	107	7	3.00	3.00	85	.00	.09
19	4*	4	113	13	.25	.46	123	.00	.00
19	4*	4	137	12	.00	1.50	130	.00	.00
							72	.00	.00
19	12	6	85	10	.00	.00	85	.00	.00
							72	.00	.00
19	12	6	97	11	.00	.00	54	.00	.00
21	5*	6	119	11	.00	.00	88	.00	.00
21	5*	6	119	10	1.25	.70	95	.00	.00
22	6	6	96	11	.00	.00	78	.00	.00
22	6	6	96	11	.00	.00	67	.00	.00
24	5*	6	103	10	1.50	.30	88r	.00	.00
24	5*	6	112	10	.00	.00	95r	.00	.00
25	1	3	111	14	.75	.43	116	.00	.00
25	1	3	91	14	.25	.43	—	—	—
26	11*	6	124	7	.00	.00	121r	.00	.00
26	11*	6	131	7	.00	.00	147r	.00	.00
28	3	6	90	10	1.50	1.20	53	.50	.16
28	3	6	95	12	.25	.16	71	.00	.00
							58r	.00	.00
35	3	6	80	5	.00	.00	72r	.00	.00
35	3	6	69	5	.00	.00	89r	.00	.00
35	6	6	77	11	.00	.00	71r	.00	.00
35	6	6	88	11	.00	.00	59	.00	.00
36	7	4	77	11	.00	.00	43	.00	.00
36	7	4	77	11	.00	.00			

* Basenjis.

† Died during experiment (all same litter).

r = repeated control.

able at all times. Once per day the mother was fed a dish of whole cow's milk prepared from powdered milk plus vitamin *A* and *D* supplement (Abbott's Haliver Malt). After weaning, the experimental puppies were fed warm Borden's Esbilac, a simulated bitch's milk, several times per day. Left-over milk was left in the room and extra feedings given if the puppies appeared hungry or not gaining weight properly. Dry food was also available as for the control puppies. No effort was made to keep the feeding on a regular schedule. In this respect the experiment is quite different from those of other authors.

The young puppies were quite clumsy in their lapping and frequently covered themselves with milk. This had to be removed by bathing, and when this happened the control puppies were treated similarly. This bathing was also necessary to make sure that any body sucking was truly non-nutritive.

The primary data were collected by daily weighing, a finger sucking test, and daily observations. This routine was begun as soon as possible after the birth of a litter. Because some litters could not be moved into the experimental room until data was complete on a previous litter, the time at which this care was started varied from the first to the tenth day of life. The finger sucking test was usually omitted on weekends. At weaning the puppies were kept apart from the mother for two weeks and then returned. Testing was continued throughout this period, except in some of the older litters in which no sucking had appeared for a week or more. In a few of the litters testing was continued after the return to the mother in order to see if there was any possible disturbance of behavior at this point.

The finger sucking test consisted of placing a wet finger in the puppy's mouth for 60 seconds. A four-point rating scale was used, ranging from no sucking to 60 seconds of strong continuous sucking. The actual rating scale was: 0, no sucking; 1, mild, non-continuous sucking; 2, either mild sucking continued for 60 seconds or strong sucking which was non-continuous; 3, strong sucking continued for 60 seconds. This test was given at 8:00 to 8:30 each morning before feeding and between 11:00 and 11:30 after the puppies had had an opportunity to eat. The score for each day is the mean of these two ratings. In the tables and graphs this has been presented for convenience as the sum of these two scores, giving a scale ranging from 0 to 6.

Behavioral observations were made at varying intervals throughout the day. Approximately one hour of observation was done daily, distributed in four 10- to 15-minute periods. The puppies were carefully watched for evi-

dence of body sucking on themselves or litter mates, as well as chewing, licking, and sucking of other objects.

C. RESULTS

1. General

An overall picture of the results is given in Figures 1 to 4. Figure 1 shows that the sucking tendencies of the puppies start with a high rating at birth and rapidly decline to the sixth day. This confirms Ross' original observation on a small sample of puppies (7). After the sixth day over half of the puppies have almost completely stopped finger sucking. During the next eight days 58 per cent of the controls and 63 per cent of the experimentals show either no finger sucking or a slight amount of sucking on one or two days.

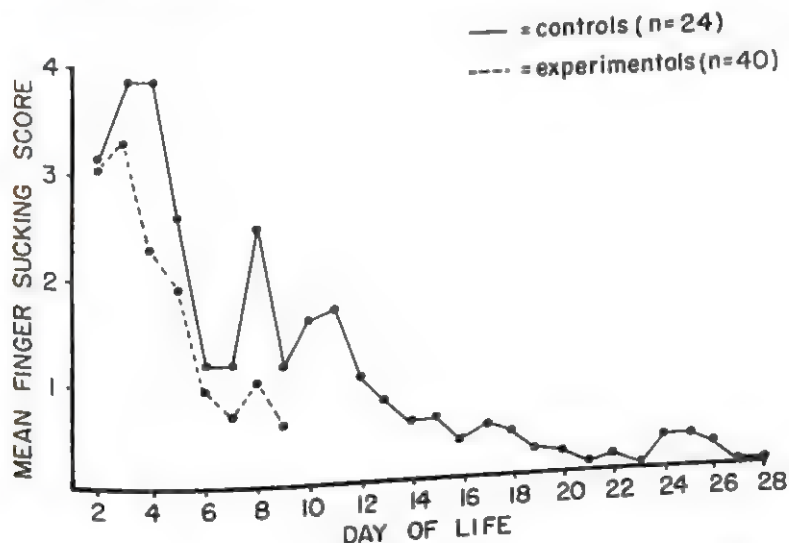


FIGURE 1

Average finger sucking scores for control puppies; and for experimental puppies prior to separation of the first animals. Note the rapid decline of sucking during the first week, and that prior to treatment the experimental puppies show slightly less sucking than controls.

There are a fair number of animals who continue sucking for a longer period. For these animals there appears to be actually a slight increase in the finger sucking tendency during the following week. Whether this is a maturational phenomenon or is caused by some other factor is not clear from the data.

Figure 1 also shows that the control animals originally sucked more than

the experimentals, so that the selection of experimental animals was actually against the hypothesis that separation should increase sucking. As will be seen later, this effect was probably produced by selecting the larger animals in the litter for the experimental subjects.

Figure 2 shows that separation has a clear-cut effect upon sucking and

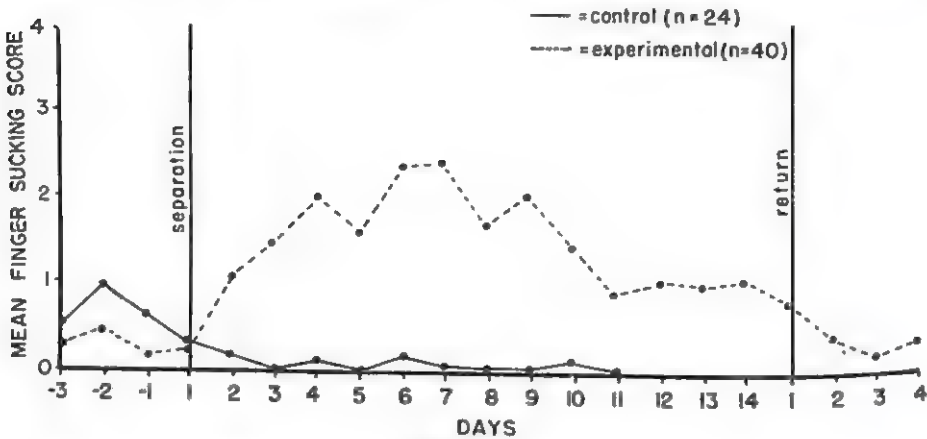


FIGURE 2

Average finger sucking scores for control and experimental puppies during separation, irrespective of age at separation. The maximum effect on the separated pups is produced 6-7 days after weaning.

that the maximum amount of sucking appears on the 6th and 7th days of separation. Figure 3 shows that there is a clear difference between puppies separated at 19 days and earlier and those separated from 20 to 28 days inclusive. Puppies weaned after 28 days of age showed no sucking whatsoever.

Figure 4 shows the relationship between weight and separation. Before separation the experimental puppies were heavier. Afterwards they tended to gain weight less rapidly than the controls and actually followed the curve formerly taken by the controls. The latter showed an upward spurt after separation, probably because of an increased available milk supply from the mother, which is the sole source of food in younger animals. This, of course, indicates that sucking is related to weight gain and, by inference, to hunger.

2. The Effect of Age

When the relative sucking scores of experimental and control puppies are graphed according to age (see Figure 5) it will be seen that there is a decline in the effect on the 19th day and following. The first experimental animal to show a zero score is found on the 19th day. Therefore the data were

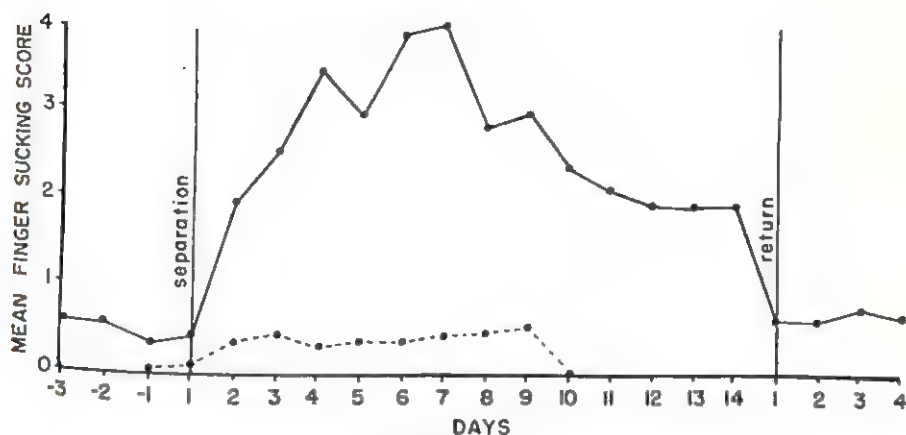


FIGURE 3

Average finger sucking scores for experimental puppies of different age groups. Solid line represents puppies weaned at 19 days and earlier; broken line, puppies weaned on Days 20-28. Puppies weaned after 28 days showed no finger sucking.

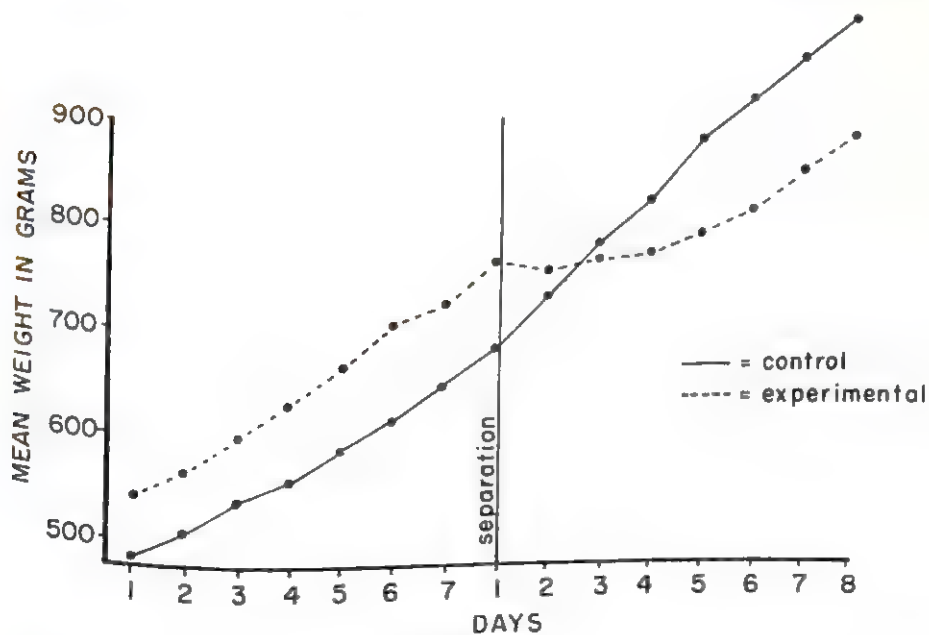


FIGURE 4

Average weight of control and experimental puppies, irrespective of age, before and after separation. Weaning lowered the weight of the separated animals, and the additional milk supply benefited their litter-mate controls.

divided at this point, leaving 18 experimental animals from 10 to 18 days of age and 16 from 19 to 28. In relation to the beginning of the critical period of socialization, one of the most reliable measures is the appearance of the startle reaction to sound. This appears on the average at 19 1/2 days, with an estimated standard deviation of 2.3 days (9). Dividing the data between 18 and 19 days therefore insures that the majority of the earlier sample had not passed into the period of socialization at the time of separation.

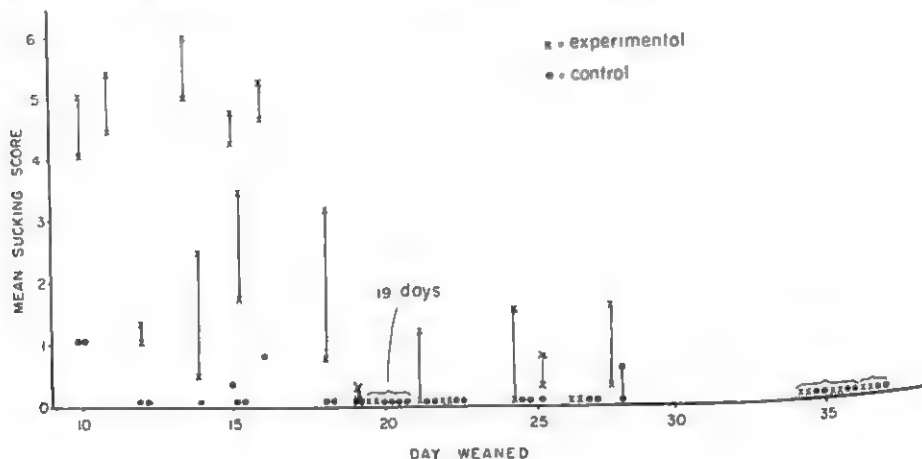


FIGURE 5

Average finger sucking scores for the four days immediately following separation. Note that experimental animals first overlap their litter-mate controls on Day 19. For precise data, see Table 1.

In order to obtain the immediate effect of age of separation the sucking scores for experimental animals were calculated for the first four days following separation. Over half of the older animals received a zero score while none of the younger animals did so. The chi-square analysis of this result gives a probability of less than .01 of obtaining this result by accident.

There remains the possibility that the age effect is due to an accidental selection of lighter animals in the young group. Nine out of 18 of the younger animals were below 90 per cent of normal weight at the last weekly birthday preceding separation, while only one out of 16 of the older animals was below this point. However, if we eliminate all animals below 90 per cent of normal weight we get nearly comparable samples in both age groups, with 55 per cent and 60 per cent respectively of animals at normal weight or above. When these two groups are compared by the Mann-Whitney *U* test the younger group shows a higher sucking score than the older, with a

probability of less than .001. The same result is obtained when the average sucking score over the entire 14-day period of separation is used. We may therefore conclude that there is a definite and important effect of age upon sucking which is independent of weight.

3. *The Effect of Weight*

As indicated above, there are indications that the amount of sucking done by a puppy is related to its weight. This phenomenon may now be analyzed in detail. In the first place, the Pearson correlation coefficient between the average sucking rate per day for all puppies from birth to the end of the experiment and the average weight gain per day for the same period is $-.55$. For the experimental puppies, the corresponding figure for the time of separation only is $-.68$. The smaller and less rapidly growing animals thus have a strong tendency to do more finger sucking than the larger ones.

This may be analyzed in relation to age and previous conditions of nutrition. The mothers differ a good deal in the amount of milk which they can produce, and this is the only source of food for puppies up to three weeks of age. In order to compare puppies of different ages, the weight on the weekly birthday preceding separation was compared to the mean weight at the same age of a larger number of other animals of the same breed reared under standard conditions. By accident of random selection, more of the animals in the early-weaned group from 10 to 18 days were under 90 per cent of normal weight than were the two older groups. If all animals from 10 to 18 days of age are used for a rank order correlation between per cent of normal weight and sucking, the result is $-.68$. If all the animals below 90 per cent are eliminated the correlation is still $-.59$. In this age group there is a strong negative correlation between weight and the amount of sucking; that is, the lighter animals suck more than the heavier ones.

If a similar analysis is made of animals between 19 and 28 days of age, there is almost no evidence of correlation, and those above 28 days of age show no finger sucking, in spite of the fact that these were animals which fell much below the normal weight. We may conclude that weight has an important relationship to sucking only in the early age group.

One more set of facts has some bearing on the analysis. The puppies were measured according to the time it took to equal or exceed their weight before separation. In many cases this took place on the first day; i.e., there was no loss of weight. In others it might be several days before the lost weight was regained. The third day was arbitrarily selected as a point indicating severe

weight loss resulting from the fact that the puppy was not able to adjust himself quickly to the new method of feeding. As shown in Table 2, nearly .8 of the youngest age group had difficulty, which was slightly more marked in the lighter puppies. Only a third of the 18 to 28 day puppies showed this difficulty, while none of the older puppies showed it.

TABLE 2
PROPORTION OF ANIMALS SHOWING SEVERE WEIGHT LOSS AFTER SEPARATION
(Lost Weight not Regained by Third Day of Separation)

Normal weight	10-18	Age 19-28	29-36
Under 90%	8/9	0/1	0/6
90% or over	6/9	5/15	0
Total	14/18	5/16	0/6

This means that in spite of every effort to keep them well fed, the puppies from 10 to 18 days of age had greater difficulty in adjusting to the new method of feeding than did the older ones. It follows that these younger animals were hungry over a longer period of time. We can conclude that the amount of finger sucking is strongly related to hunger in young puppies of 18 days of age and younger. Hungry puppies of this age tend to suck on any available object. Beyond this age the puppies adjust themselves more rapidly to the new method of feeding and hence become less hungry. However, since there was no correlation in this group between weight and sucking, we can conclude that there is no longer a strong tendency toward non-nutritional sucking in response to hunger.

4. *Supplementary Experiment on Effects of Hunger*

A further experiment was done in order to test the effect of hunger as separated from the possible emotional disturbances of prolonged separation from the mother. In a litter of five cocker spaniels, puppies were separated from the mother on alternate days, so that there were always at least two puppies with the mother and no puppy was exposed to prolonged food deprivation. During the first week of life the puppies were separated for periods of four hours. During the second week they were separated for eight hours, and during the third and fourth weeks they were separated 16 hours. The latter procedure was carried on at night instead of the day. The finger sucking test was given before and after the period of brief separation.

The results showed significant differences in finger sucking and other oral activities only during the third week, the age period in which significant

results were given by the experiment on sudden weaning. It may be concluded that food deprivation causes an increase in sucking up to the end of the third week, but not thereafter.

The negative results prior to 14 days do not necessarily indicate that before this time there is no effect of hunger on sucking. During the first week the ratings on the control tests were so high that no additional effect could have been obtained by hunger. There was also a tendency in this litter for sucking activity to show a daily rhythm, with higher amounts of sucking in the late afternoon than in the morning. The scores for both experimental and control subjects tended to rise in the afternoon. If the 8-hour deprivation during the second week had been done at night, it might have given some results during the second week. Also, the period of deprivation was considerably shorter than that given later and might not have been effective.

The behavior of the hungry puppies in free situations showed some differences from week to week. During the first week the puppies showed considerable random exploratory activity and were frequently found in different parts of the room. During the second week the puppies were found huddled together in the spot in which they had been originally placed, apparently showing more responsiveness to each other. In the third week the puppies tended to locate themselves at the point in the fence nearest to the location of the control puppies. This indicates that the eyes (now open) were being used for visual localization. Finally, during the fourth week, the puppies were much more active and tended to come to the fence and yelp. All these changes are consistent with the regular pattern of maturation of sensory and motor capacities.

This experiment confirms the conclusion of the previous section that hunger increases the amount of finger sucking in animals 18 days of age or younger but not in older animals.

5. *Body Sucking*

One of the striking results of the present experiment was that no body sucking was observed between the pairs of experimental puppies. This has, however, occurred in certain other experiments and can be described as follows. Unlike the human infant, the dog has no convenient part of its own body which can be sucked. The paws are covered with hair and cannot be gotten easily into the mouth. Animals which show this type of non-nutritional sucking are most likely to suck the under parts of the bodies of other animals, either the penis of males or the vulva of females, and occa-

sionally the tip of the tail. This seems to last for only a short period in the life of puppies, probably because as they grow older they have a tendency to fight off any puppy that nurses upon them.

In contrast to the experiment described in this paper, J. A. King and O. Elliot observed several cases of body sucking in an experiment involving sudden weaning. Four litters of basenjis and one litter of fox terriers were weaned at approximately four weeks of age (27-30 days). These 16 animals were divided into groups of four and individually fed twice per day on a regular schedule. Two groups were given an excess of food and two a limited amount, with the result that the weights began to rapidly draw apart. The group given an excess of food had dry food available in the pen. They ate very little of this and gorged themselves on the scheduled feeding of a mixture of milk and mash.

Body sucking was observed in all four groups, at ages ranging from five to seven weeks (38-50 days). Body sucking was observed both before and after feeding. It occurred in four out of 11 basenjis and two out of five terriers. The two terriers were the smaller animals in the litter, both at weaning and at the time when body sucking was observed, but the basenjis which showed sucking were intermediate in size. The basenji group was unusual in that it included several unusually small litters, three litters of two and one of five, while the fox terriers came from a litter of five. Five of the cases of body sucking came from the two large litters (10 animals in all) while only one came from the small litters (six animals).

The numbers are too small to determine why only part of the animals developed body sucking. Underfeeding can be eliminated, since sucking occurred in the overfed litters. Body size may be involved but gives no significant results with the numbers used. The data suggest the possibility that animals in the two large litters were subjected to frustration previous to weaning, either through lack of opportunity to nurse, or shortage of milk, or both. However, the numbers are again too small to draw a definite conclusion.

The important and clearcut result of this experiment is that the phenomenon of body sucking can be produced by sudden weaning followed by regular, scheduled feeding. This type of feeding apparently has the effect of making the animals intensely hungry at particular times of day.

During a 10-year period, careful daily observations of behavior have been kept on some 500 different puppies raised in another experiment in which animals were left with the mother until 10 weeks of age, which is beyond

the normal weaning time. Beginning at three weeks the puppies were given one daily scheduled feeding of milk. Dry mash was available at all times and the puppies began eating it sometime after three weeks of age. Body sucking was observed in only one litter, following early weaning caused by the death of the mother.

This was a litter of F_1 hybrids with a basenji father and cocker spaniel mother. The litter developed in a reasonably normal fashion through the first two weeks. At 17 days of age it was noted that the mother was not allowing the puppies to nurse freely, and that they appeared hungry. The milk supply of the mother seemed to be failing, and powdered solid food was made available for the puppies, which were able to take some of it. The mother continued to let the puppies nurse occasionally for the next 10 days. On the 28th day after the birth of the puppies she became acutely ill and died. At this point the puppies had to depend on the scheduled daily feeding of milk for their entire milk supply. Judging from the previous slow gain in weight they had been quite hungry for at least a week before death of the mother.

Two kinds of frustration were observed in this litter. The mother did not allow the puppies to nurse as long as they wished, and they also were unable to obtain sufficient food to relieve their hunger.

The puppies were first observed sucking on each other's bodies when they were 35 days of age, which is very close to the time when this was first observed in the experiment cited directly above. These puppies would lie in a group, sucking each other's tails, ears, penes, and vulvae. This activity gradually disappeared in the next few weeks as the puppies grew older.

The two cases where body sucking occurred, one experimental and one accidental, have two features in common. One is that the puppies were completely weaned at four weeks of age, and the other is that the body sucking appeared approximately one week later while the puppies were given food in a scheduled fashion. The major experiment described in this paper was unlike these cases in that feeding was not done in a scheduled fashion and the puppies therefore did not become acutely hungry at particular times of day. The puppies were kept away from the mother for a period of two weeks so that there should have been opportunity for body sucking to develop.

The data in this section suggest that, in addition to the hunger produced by scheduled feeding, earlier feeding frustration and/or sucking frustration are also involved.

These observations may be compared with those of a second litter of hybrid puppies with a basenji mother, in which body sucking did not occur. This female developed a case of diarrhea which did not respond to treatment until the puppies were between four and five weeks of age. Up to this time the milk supply of the mother was deficient, and the puppies were much retarded in their growth. Afterwards the mother's supply appeared to be restored to normal. Unlike the cocker mother, the basenji allowed her pups to nurse. While they showed considerable evidence of suffering from hunger, they were never seen to suck on each others' bodies, either during the period of frustration or the normal period thereafter. Unlike the previous cases, these puppies were not subjected to violent hunger as they grew older, either because of periodic feeding or a short food supply.

D. DISCUSSION

1. *Comparison with Other Experiments on Puppies*

The most elaborate and detailed experiment of this sort was done by Levy (4), with a litter of puppies which were given a great deal of individual care and attention and on which detailed measurements and observations were made over a long period. The experiment was similar to the present one in that hunger was a factor. At the end of the experiment, at 30 days of age, the control puppies left with the mother averaged three pounds heavier than those used as experimental animals, although the latter had gained approximately two pounds each over a three-week period.

In Levy's experiment four out of six puppies were weaned at 10 days of age. Nothing can be said about the previous state of nutrition since we have no data for normally raised puppies in the collie breeds. The treatment after weaning was to raise the animals on a bottle on regular scheduled feedings starting three hours apart, beginning at 7 A.M. and ending at 11 P.M. The care given was very similar to that given to human infants. The experimental puppies were kept in pairs in boxes up to the 19th day of life, after which they were separated except at feeding and testing time. The chief experimental variable was to feed one pair of puppies with bottles having small holes in the nipples, so that it took much more sucking to elicit the same amount of food than it did for the other pair who were fed with large-holed nipples.

Levy's results are reported in great detail. He observed some body sucking in both groups of animals, but it was much more frequent in the rapid-feeding pair. Body sucking was observed as early as the 14th day, occurring

both before and after the 19th day. It is possible that keeping the animals in boxes instead of a large room as in the present experiment gave more opportunity for body sucking. This would be particularly important at the early age when puppies could get separated.

Levy's finger sucking test was not begun until the 15th day of life. From the 15th to the 19th day he obtained more finger sucking in the slow feeders. After the 19th day there was much more finger sucking in the rapid feeders. Levy interprets his results in terms of the frustration of an impulse to suck, so that animals deprived of nutritional sucking tend to do more non-nutritional sucking.

Fleischl (3) repeated Levy's experiment with three mongrel dogs, two short-time feeders and one long-time feeder. The puppies were kept in a home, and their house was a dresser drawer, so that they were always in close contact. An attempt was made to raise them as much as possible like human triplets. No detailed results were given, but Fleischl reports that the rapid feeders showed more intense finger sucking throughout the experiment, in contrast to Levy's finding that slow feeders sucked more at an early age.

Lindzey (5), in a pilot experiment, attempted to reproduce Levy's experiment, starting with a litter of beagles at 10 days of age. Because of various difficulties the puppies were returned to the mother at 16 days of age. No body sucking was observed at later ages.

In summary, the puppy experiments are difficult to compare because of the different techniques employed. All of them are alike in that the weaned animals were not as well fed as those kept with the mothers. Body sucking has been produced only in animals which have been fed on a regular schedule after weaning. It has been produced in late-weaned pups, when the pups are weaned to the cup or solid food (King and Elliot, etc.), and in early-weaned pups which were weaned to a bottle (Levy).

All these experiments included the factor of hunger and some evidence of sucking frustration, and it is not clear whether either of these factors or both are necessary to produce the effect.

The animal experiments lead to the conclusion that the phenomenon of non-nutritional sucking is not a simple one based on a single causal factor or situation. We can conclude that there are at least five possible factors which should be considered in any future set of experiments. The first of these is genetics. In normally raised litters of puppies there are some animals which tend to be suckers and some which do not. While we have no

experimental proof for this factor, it is likely that both breed and individual differences exist in the tendency to suck non-nutritional objects. Besides this, there are differences in the care and nutrition provided by the mother before weaning.

Second, the age at weaning is an important factor, as shown by the present experiment. Puppies weaned at 19 days and after show quite different sucking responses from those weaned earlier, and animals weaned after 28 days of age show different responses again.

Third, there is the experience of the animal prior to weaning. Part of this experience is the varying amount of nutrition received from different mothers, leaving the puppies in different nutritional states. The other part of this experience is the degree of frustration which the puppy receives from the mother. Frustration due to hunger apparently has no effect by itself, but does in combination with frustration of sucking. Frustration of sucking without hunger has not been attempted in any experiment.

The fourth variable is the treatment after weaning. Weaning to a bottle apparently produces a different effect from weaning to a cup, and weaning to regular scheduled feeding a different effect from weaning to demand feeding.

Finally, there is the factor of hunger, which has been present in all of the experiments. The question of whether sucking frustration in the absence of frustration connected with hunger would produce an effect has not yet been experimentally tested. Final and definitive results on this subject will depend on an adequate program of experimental work, using large numbers of subjects and controlling for at least the five factors outlined above.

2. *Comparison with Human Data*

Any comparison of the results of studies of human children with the present experimental work on puppies depends upon an accurate comparison of developmental stages. Human neonates do not have the major sensory deficiencies of puppies. In humans, both eyes and ears are responsive to stimulation at birth or shortly afterwards. On the other hand, puppies are more advanced in motor development, being able to crawl at birth and making later advances much more rapidly.

Both sorts of infants normally obtain their first nourishment by nursing. It is difficult to compare the natural times for changing over to other kinds of food, since human infants will assimilate a variety of specially prepared foods almost from birth. Under more primitive conditions it would appear

that no solid food could be managed by the infant until the development of the first set of teeth. The first incisors appear at about six months of age, and the complete set of deciduous teeth is usually present by one and a half to two years of age. This time may be taken to roughly correspond to the period beginning at three weeks in the puppy, when the first teeth appear. Newborn human infants can be fed successfully by a cup with careful manipulation. In this respect newborn babies resemble puppies in the transition stage (two to three weeks), which are able to eat by lapping with some assistance.

The development of human learning capacities has been reviewed by Munn (6), who comes to the conclusion that some unstable conditioning can be obtained at birth, and possibly even prenatally. Good evidence of stable conditioning does not appear until approximately one month of age. The further development of human psychological capacities is perhaps best indicated by the work of Spitz and Wolf (10) on the smiling response, which is aroused by any human face or mask at one month of age. Discrimination between familiar and unfamiliar faces does not begin until five or six months. It may be concluded that in learning capacity the human infant from birth until one month of age is roughly equivalent to that of the puppy in the transition stage from two to three weeks. Both show unstable conditioning. From one to six months the human infant is equivalent in learning capacities to the puppy from approximately three to four weeks, in the early part of the period of socialization.

In the human experimental studies it is generally assumed that the infants are well fed, and all are fed on regular schedules. Most experiments have been done on extremely short periods of development and are frequently compared with those done at widely different ages. Two highly significant experiments have been done on babies during the first 10 days of life. Davis *et al.* (2) fed newborn babies by the bottle, cup, and breast methods and tested their response to a finger sucking test. They found that the breast fed babies showed more finger sucking than the others. Like the newborn puppies, there was some tendency towards a decline as they got older. Figures are not available on the nutrition of these infants, but it is likely that breast fed infants obtained less nourishment than the other two groups. The authors interpret the experiment as showing the results of learning, namely that the activity which is rewarding tends to be continued. This would also imply that the babies had no power of discrimination at this age.

Brodbeck (1) has reported a second experiment which more closely

parallels that described in this paper. He bottlefed a group of newborn infants for four days, and then changed them suddenly to cup feeding for another four days. This would be the equivalent of sudden weaning in the transition stage of the puppy. The chief difference is that the human infants were not allowed to become hungry. The bottle fed babies were treated in various ways, some getting less milk per feeding and others getting a more rapid flow of milk. All of the infants showed less tendency to suck the fingers when changed to cup feeding. The possibility still remains, of course, that the cup fed infants are better fed than the bottle fed ones and that the decline may be due to decreasing hunger. It may be concluded, however, that sudden weaning at this stage in development, when not accompanied by hunger, does not produce a tendency toward thumb sucking.

It should be pointed out that neither of these two experiments produced the persistent type of non-nutritional sucking which forms a behavior problem in older infants and which is comparable to the body sucking produced in Levy's puppies and those of King and Elliot. This early sucking of human infants may correspond to the transitory non-nutritional sucking seen in newborn puppies.

In conclusion, it is clear that there are two kinds of sucking behavior in puppies. One of these is typical of newborn puppies and consists of sucking anything resembling a nipple which is thrust into the mouth. Such behavior normally disappears during the first week, presumably because the puppies become well fed, but it can be brought back at later ages by making the animal hungry. It is difficult to elicit after 18 days of age, presumably because the puppy is now able to discriminate between nutritive and non-nutritive objects, and it disappears entirely after 28 days of age. This kind of sucking is therefore strongly influenced by hunger and age.

The second kind of sucking is the deliberate non-nutritional sucking on special objects by older animals. This is usually directed toward the bodies of litter mates. The situation which produces it is some sort of rigid feeding schedule. This in turn produces acute hunger at special times, with the result that the animal has a tendency to eat very rapidly. As Levy showed, this kind of sucking can be greatly reduced by forcing the animals to eat more slowly, using (in his experiment) a bottle with a small-holed nipple. In all experiments in which this type of sucking has occurred there has been evidence of frustration of both eating and sucking after the age of 18 days. Sometimes these factors have been part of the experiment and sometimes they have been accidental.

The effect of age has been thoroughly explored in the present experiment in connection with the first type of sucking. With this data and other experimental results it should now be possible to do a thorough experiment with the effect of age on the genesis of the second type of sucking.

These two types of sucking have not been hitherto distinguished, either in animal experiments or in human subjects, but they are clearly and recognizably distinct. It is probable that the sucking observed in young babies in the first 10 days of life belongs to the first type, which is, of course, an adaptation for biological survival. The second type of sucking can be identified with the problem thumb sucking of older children. It is to be hoped that in the future methods of clear diagnosis can be devised. In the case of human infants the two kinds can be easily confused because the young baby is likely to bring its thumb into contact with the mouth in the course of random movements. The recognition and diagnosis of the adaptive sucking of young infants apart from the mildly neurotic form seen in older ones may help to eliminate some of the confusion which now exists in the literature (8).

If the same factors are assumed to develop thumb sucking as appear to produce body sucking in the dog, there are two things which may be helpful in avoiding it: (*a*) allowing the child to nurse or suck as long as it wishes, as suggested by Levy; (*b*) relaxing a rigid feeding schedule to permit demand or ad lib feeding. There may also be a critical age for the production of this behavior, probably a considerable time after birth. However, these possibilities should be tested by experiment before they are accepted.

E. SUMMARY

1. A total of 14 litters and 64 subjects were used in the experiment. Three litters were African basenjis; the rest cocker spaniels.
2. Pairs of puppies were suddenly weaned at ages ranging from 10 to 36 days. Forty experimental and 24 control animals were studied.
3. A daily finger-sucking test was given before, during, and after the weaning period. Additional data were collected through weights and daily observations.
4. Normally reared puppies show a high rate of finger sucking at birth, with a rapid decline through the 6th day. About 60 per cent show no sucking after the first week; the rest show slight amounts of sucking for varying periods thereafter.
5. Eighteen animals weaned from 10-18 days of age all showed more sucking than litter mate controls; seven out of 16 animals weaned from

19-28 days showed more sucking than controls; the six animals weaned after 28 days showed no sucking.

6. Puppies weaned before 19 days show significantly more sucking than those weaned from 19-28 days, even when the samples are adjusted for differences in previous state of nutrition.

7. Animals weaned from 10-18 days show an *RO* correlation between sucking scores and per cent of normal weight at weaning of -68 . There is no evidence of this relationship in animals weaned at 19-28 days.

8. It is concluded that sudden weaning produces an important increase in non-nutritive sucking in animals under 19 days of age, and that this increase is inversely related to the previous state of nutrition. This effect is greatly lessened in older animals, where it is not related to previous state of nutrition.

9. This change in behavior coincides with the beginning of developmental period III, Socialization. It also coincides with the first appearance of the capacity for conditioning and hence probably the first appearance of discrimination learning.

10. Persistent body sucking was not observed in any of the above group, but it did occur in six out of 16 animals in another experiment involving sudden weaning. In this latter experiment animals were fed twice per day on a regular schedule instead of ad lib as in the former experiment.

11. A comparison of experiments in which persistent body sucking occurs shows that two conditions are necessary: (a) a periodically high degree of food motivation or hunger produced by a regular feeding schedule, and (b) a lack of opportunity for normal nursing. Other factors which are probably important are heredity, age, state of nutrition, and previous nursing experience.

12. It is concluded that there are two kinds of sucking behavior in puppies. The first is the tendency of newborn animals to suck anything resembling a nipple, if thrust into the mouth. This is increased by sudden weaning and aggravated by insufficient nourishment, but is difficult to elicit after 18 days of age. The second type is deliberate non-nutritive body sucking by older animals, associated with rigid feeding schedules and the curtailment or absence of normal nursing.

13. It is suggested that these types of non-nutritive sucking can also be differentiated in human infants; i.e., adaptive sucking of young infants and the mildly neurotic thumbsucking of older infants.

14. It is suggested that avoidance of a rigid feeding schedule may be

helpful in preventing sucking of the latter type. This is in addition to the method suggested by other authors, of providing adequate opportunity for sucking.

REFERENCES

1. BRODBECK, A. J. The effect of three feeding variables on the non-nutritive sucking of new-born infants. *Amer. Psychol.*, 1950, **5**, 292-293.
2. DAVIS, H. V., SEARS, R. R., MILLER, H. C., & BRODBECK, A. J. Effects of cup, bottle, and breast-feeding on oral activities of new-born infants. *Pediatrics*, 1948, **3**, 549-558.
3. FLEISCHL, M. F. The problem of sucking. *Amer. J. Psychother.*, 1957, **11**, 86-97.
4. LEVY, D. M. Experiments on the sucking reflex and social behavior of dogs. *Amer. J. Orthopsychiat.*, 1934, **4**, 203-224.
5. LINDZEY, G. An experiment on sucking frustration in young puppies. (Unpublished note, Roscoe B. Jackson Memorial Laboratory.)
6. MUNN, N. L. Learning in children. Ch. 7 in *Manual of Child Psychology* (ed. by L. Carmichael, 2nd ed.). New York: Wiley, 1954.
7. ROSS, S. Sucking behavior in neonate dogs. *J. Abn. & Soc. Psychol.*, 1951, **46**, 142-149.
8. ROSS, S., FISHER, A. E., & KING, D. Sucking behavior: A review of the literature. *J. Genet. Psychol.*, 1957, **91**, 63-81.
9. SCOTT, J. P. Critical periods in the development of social behavior in puppies. *Psychosomat. Med.*, 1958, **20**, 42-54.
10. SPITZ, R. A., & WOLF, K. M. The smiling response, a contribution to the ontogenesis of social relations. *Genet. Psychol. Monog.*, 1946, **34**, 57-125.
11. WHITING, J. W. M. The cross-cultural method. Ch. 14 in *Handbook of Social Psychology* (ed. by G. Lindzey). Cambridge: Addison-Wesley, 1954.

Roscoe B. Jackson Memorial Laboratory
P.O. Box 847
Bar Harbor, Maine

EFFECTS OF DIFFERENTIAL INSTRUCTIONS AND ANXIETY LEVEL ON DISCRIMINATION LEARNING*

Madigan Army Hospital, Tacoma, Washington

ROBERT E. MINNICH AND PERRY LONDON

A. PROBLEM AND HYPOTHESES

Research on the relationship between scale-defined anxiety and various aspects of the learning process has rather consistently shown that anxious individuals are prone to overgeneralize threatening stimuli in their environments and are less able than others to discriminate stimuli (1, 3, 10, 12, 13). It has been found too, that experimentally induced stress conditions tend to alter some learning performances of high-anxious subjects, but the direction of this alteration is variously reported as a disruption or facilitation of the learning process. Taylor (13), for example, alleges that test-anxiety is a measure of an individual's "readiness to respond," which might in some situations hasten learning, in others impede it.

A recent study by London (7), investigating the relationship between generalization and adjustment, found that maladjusted children of all ages tended to systematically overgeneralize stimuli relative to normals, and concluded that maladaptive behavior is characterized by faulty discrimination in the perception of relatively innocuous stimuli. The instrument employed to measure generalization-discrimination was a series of graded picture tests, similar to those used in concept formation studies by Frenkel-Brunswick, Zaslow, and others.

Since there is generally a positive relationship between maladjustment and anxiety, it is likely that children with high anxiety, like the overtly maladjusted, overgeneralize responses to *innocuous* as well as to threatening stimulation, this tendency becoming more apparent under stress than under neutral conditions. The present study investigated this general hypothesis.

One important criticism levelled at previous studies of this kind is that experimentally induced stress conditions, such as electric shocks, were not related to the manifestations of anxiety on which selection of subjects depended. In order to better control this relationship, the present study em-

* Received in the Editorial Office on July 23, 1958.

ployed a scale measuring anxiety about school examinations as the operational criterion for anxiety and the facade of a school examination as the operational stress condition, thus making the source of stress specific to the source of anxiety. The following operational hypotheses were therefore established:

1. Children who score high in test anxiety differ from children scoring low in test anxiety in ability to discriminate graded stimuli.
2. Differences in discrimination between high and low test anxious subjects are larger under test (stress) conditions than under nontest (neutral) conditions.

B. METHOD

1. *Materials*

Materials for the discrimination tasks consisted of five of the six experimental sets of pictures employed by London (7). Three of the sets, representing "social" objects, were: (a) an ogre holding a club, changing to Santa Claus, holding a bag of toys; (b) a young man in a suit, changing to a young woman in a suit; and (c) a little girl standing at a table, changing to a cocker spaniel standing at a table. Two of the sets, representing "Non-social" or "abstract" objects, were: (a) a triangle which changed to a circle; and (b) a true blue square, which changed to a true green square.

A training, or practice set, was administered to the subjects prior to the test sets. The practice set consisted of a gun (rifle), which changed to a broom over a series of five cards, as opposed to fifteen cards in each of the test sets.

2. *Subjects*

In order to eliminate possible variability associated with sex differences, the final sample was limited to boys. Sixty boys were selected from grades three through five of an elementary school; 16 were from the 3rd, 28 from the 4th, and 16 from the 5th grades. They ranged in age from 100 to 146 months. S's were divided equally into four groups which were matched for grade, age, *IQ*, and socio-economic status. Their selection and assignment to groups is described below.

3. *Procedure*

The Sarason Test Anxiety Questionnaire, Children's Form, was employed to measure anxiety levels. It was administered orally to 355 children, including 165 boys, in their respective classrooms, while the teacher was not in the room. S's were informed that there were no right or wrong answers and that no grownups would see their answers.

The anxiety score for each child consisted of the total number of Yes's circled on his answer sheet. The higher the score, the greater the test anxiety. A score at or above the 80th percentile of each grade was chosen to represent high test anxiety, and a score at or below the 20th percentile represented low test anxiety. Mean scores were 10.5, 12.8, and 15.1, for 3rd, 4th, and 5th graders respectively.

The positive relationship indicated between grade and test anxiety is in agreement with Sarason's findings (12). This relationship necessitated that a distribution of anxiety scores be made for each grade separately rather than for all grades combined. Matchings for age, *IQ*, and socio-economic status were similarly made by grade level. Of the children initially chosen on the basis of their anxiety scores therefore, an equal number from each grade were randomly assigned to alternate experimental groups. Composition of final groups in terms of age and *IQ* (Otis Quick Scoring *MAT*, New Ed., Form *A*) is presented in Table 1; analysis of variance reveals no significant differences. Socio-economic status, as indicated by parental occupation, was Middle Class in all cases.

Thirty high test anxious (*HA*) and 30 low test anxious (*LA*) subjects were selected on the basis of the above criteria. Both high- and low- anxiety groups were then divided, for the administration of the picture tests, into two respective sub-groups of 15 each. One group each of *HA* and *LA* subjects received "test" directions (*TD*), and the other two groups received "neutral" directions (*ND*).

TABLE 1
MEAN *IQ*'S AND AGES (IN MONTHS) FOR FINAL EXPERIMENTAL GROUPS
(*N* = 15)

Experimental group	<i>IQ</i>	Age
<i>LA,TD</i>		
Mean	103.40	122.87
<i>SD</i>	13.71	11.62
<i>LA,ND</i>		
Mean	104.80	121.13
<i>SD</i>	8.46	9.97
<i>HA,TD</i>		
Mean	97.97	121.73
<i>SD</i>	10.25	12.46
<i>HA,ND</i>		
Mean	103.00	120.47
<i>SD</i>	10.10	9.91

Subjects were tested individually, under daylight conditions, by an *E* with whom they were unacquainted. After preliminary remarks, the following instructions were given:

Test Directions (TD) groups: . . . I am going to give you a test, with five different parts. I want you to do as well as you can. I want to see if you can do as well as the other boys who take this test. Try your very best.

See this picture? (*E* held up Card 1 of the practice set.) What is it? Yes, it's a gun (rifle). Now this gun turns into something else. You are to tell me what it turns into. All right, what is this next picture?

E showed Card 2 of the practice set and the succeeding cards in order, asking "what is it?" until *S* stated a change of percept appropriate to the card. *E* then said:

That's pretty good. You finally guessed it right. It turned into a ——— (broom, etc.). That was an easy one for practice. Now I am going to show you some other pictures which also turn into something. Each time you are to tell me what it is that you think it turns into. Now for the real test.

E then proceeded with Card 1 of the first test set, and each succeeding card in order, the query "what is this?" accompanying each presentation of each card. When two consecutive "appropriate" reports of changed percepts were elicited, the number of the card evoking the first changed percept was recorded as the score for that task. All five tasks were presented in the same fashion.

The general testing procedure was identical for the "Neutral Directions" (*ND*) group, but the instructions preceding administration of the test sets differed as follows:

. . . I am going to show you some pictures I have drawn and I think you will like them. I have made up a game with these pictures which I think is fun to play. I want to see if children like it. I'll show you how it goes.

See this picture? What is it? . . . (The practice set is then administered as to the *TD* groups.) *E* then said: That's the idea, it turned into a ——— (broom, etc.). Do you understand how this game works? Now I am going to show you some other pictures which also turn into something. Each time you are to tell me what it is that you think it turns into. Here we go.

The sequence of the five tasks was randomized among the subjects.

Percept changes were considered appropriately scoreable when a child's response to a card no longer contained any reference to the initial percept.

If, for example, on the Ogre-to-Santa Claus series, a subject reported "cave man" to Card 5, "friendly cave man" to Card 6, and "Santa" to Cards 7 and 8, his score was recorded as 7. It was possible to obtain a raw score on any one task of from two to 15, so that a low score represented an operationally more rapid discrimination.

Three scores were computed for each subject: An Over-all score, representing the average score for all five tasks; a Social score, representing the average score for the three social tasks; and a Nonsocial score, representing the average score for the two nonsocial tasks.

C. RESULTS

Discriminations on the Social tests tended to appear in the middle of the card series for all groups, while discriminations on the Nonsocial tests were generally delayed somewhat beyond the middle of the series. Table 2 shows the means and variances of all four groups for the three discrimination scores.

TABLE 2
MEANS AND VARIANCES OF PICTORIAL DISCRIMINATION TEST
SCORES FOR ALL EXPERIMENTAL GROUPS
($N = 15$)

Experimental group	Combined	Social	Non-social
<i>LA,TD</i>			
Mean	9.213	8.555	10.230
Variance	.231	.311	.889
<i>LA,ND</i>			
Mean	9.493	8.800	10.770
Variance	.359	.673	1.531
<i>HA,TD</i>			
Mean	8.987	7.866	10.670
Variance	.311	.674	.702
<i>HA,ND</i>			
Mean	9.133	8.511	10.030
Variance	.215	.654	1.695

Trends of the mean values are identical for the over-all and Social tests. In each case, the High Anxiety groups, regardless of Test Directions, made a slight but reliably more rapid discrimination than the Low Anxiety groups. The *HA,TD* Test Directions group made the most rapid discriminations of the four experimental groups on both the Over-all and Social tests, while the *LA,ND* group made the slowest discriminations on all the tests (Table 3).

The trend of the means on the Nonsocial tests was slightly different, in

that both the *LA, TD* and *HA, ND* groups discriminated more rapidly than the *HA, TD* group; these differences are not, however, dependable (Table 3).

TABLE 3
"t" VALUES (2-TAILED TESTS) FOR COMPARISONS OF GROUP
MEANS ON DISCRIMINATION TESTS

Test scores	<i>LA, TD</i>	Experimental group <i>HA, TD</i>	<i>LA, ND</i>
Over-all			
<i>HA, TD</i>	1.56		
<i>LA, ND</i>	1.41	2.40*	
<i>HA, ND</i>	.47	.78	1.84#
Social			
<i>HA, TD</i>	2.10*		
<i>LA, ND</i>	.75	3.12**	
<i>HA, ND</i>	.13	2.17*	.98
Nonsocial			
<i>HA, TD</i>	1.35##		
<i>LA, ND</i>	1.35##	.26	
<i>HA, ND</i>	.48	1.60##	1.61##

$p < .20$.

$p < .10$.

* $p < .05$.

** $p < .01$.

In general, anxiety level appears to be a primary influence on rapid discrimination and test directions a secondary one. They appear moreover to be independent, again with the exception of the Nonsocial tests (Table 4).

TABLE 4
INTERACTIONS BETWEEN TEST DIRECTIONS AND TEST ANXIETY ON
PICTURE DISCRIMINATION TESTS
($N = 1$)

Source	Mean square	F
<i>Over-all Tests</i>		
Directions	.683	
Anxiety Level	1.290	2.44#
<i>DxA</i> Interaction	1.055	4.61*
<i>Social Tests</i>		
Directions	2.964	
Anxiety Level	3.584	4.04*
<i>DxA</i> Interaction	.601	4.88*
<i>Nonsocial Tests</i>		
Directions	.04	
Anxiety Level	.34	
<i>DxA</i> Interaction	5.11	4.26*

$p < .10$.

* $p < .05$.

D. DISCUSSION

Discrimination scores for the Social and Over-all tests suggest that test anxiety and test directions relate to discrimination in much the same way, i.e., by facilitating it. The effects of the two variables therefore, may well be cumulative, as suggested by the distributions of scores for the four groups. Both Anxiety groups responded to the stress of test instructions in the same manner on Social and Over-all tests, although the difference was greater for the *HA* groups. The finding that high test anxious subjects tended to respond more readily to the "anxiety cues" provided by the test directions may be evidence for their "greater potential for anxiety arousal," while differences in scores between *Ld* and *HA* groups irrespective of Test Directions may reflect, as Taylor suggests, differences in "chronic emotional states." In view of the different performances elicited by the Nonsocial tests, however, as well as those unidentified cues to which groups may have responded, this interpretation seems inconclusive.

Operationally, the *HA* and *TD* groups did appear to have been alerted to report a percept change more rapidly. Hoch and Zubin would relate this behavior to anxiety, suggesting that "a perception of threat represents the alerting effects of lesser degrees of anxiety" (4). Test situations are probably perceived as more threatening than others, and the individual is hence alerted prior to taking the test. In this context, differences in "test anxiety" may represent differences in tendencies to respond alertly specifically in test situations.

If perception of threat is an initial motivating stimulus, then more rapid discriminations may also reflect an individual's desire to avoid or withdraw from an uncomfortable situation. His alertness, stimulated by the threat, will then be focussed on making responses which enable him to withdraw from the situation, accomplished in this case by reporting an appropriately changed percept as soon as possible.

Another factor which may have entered into rapid performance on the discrimination tests is the need for social approbation; this need may be experienced particularly strongly in relation to school and school performance, and the Test Directions emphasized invidious or approval-seeking motives in encouraging children to "do as well as the other boys." Alertness, anxiety, avoidance, and need for approbation are thus all concepts which may be involved in the facilitation of discrimination performance.

Performances on the Nonsocial tests, however, take exception to the hierarchy of scores. Children in all groups discriminated consistently more slowly

here than on the Social tests. More important, however, is the fact that the performance of *LA* children improved, as on the Social tests, under the stress of test directions, while that of *HA* children, contrarily, was poorer under those conditions.

Sarason and Mandler (11), who found a similarly debilitating relationship between anxiety and stress, offer the interpretation that previous learning in test situations mobilizes a large number of irrelevant response tendencies in anxious subjects, and that these tendencies lower their performance under stress. Applied to the Nonsocial tests in this study, one could then say that *LA* subjects responded to the alerting effects of Test Directions without being hindered by attendant increases in irrelevant responses.

The consistently different performances of the *HA* subjects on Social and Nonsocial tests under different stress conditions, in terms of Sarason and Mandler's hypothesis, might be accounted for in terms of the different properties of these tests. The greater complexity of figures on the Social tests may have actually made the presence of a large number of potential responses an asset to discrimination, since any of several responses was acceptable as a final percept; the range of potentially correct percepts on the Nonsocial cards was relatively limited, thus making the presence of competing responses a hindrance. The complexity of the Social Cards, in other words, may have invited a variety of intermediate response hypotheses, offering the opportunity for the continuous reinforcement of potentially correct responses and the weakening of incorrect responses, and thus creating a context for the *HA* children which simultaneously increased alertness and reduced threat.

Whether or not anxious children have a relatively greater fund of available responses in a given situation, however, it is apparent from the data that their behavior reflects a differential sensitiveness to certain classes of stimuli relative to less anxious children. The interaction of high anxiety and few cues resulted in relative overgeneralization of stimuli, while the interaction of high anxiety and numerous cues resulted, for these children, in rapid discrimination behavior. Studies which report both rapid discrimination and overgeneralization behavior by subjects responding to the same set of cues may, therefore, be neglecting the importance of differences in internal, or mediated, cues as they relate to discrimination behavior from subject to subject. Selecting subjects on the basis of their responses to a questionnaire which samples behavior related to a specific class of cues, and presenting different types of percepts varying widely in number of cues available for

reporting changes, seems to have operated to prevent individual differences in the number of internal cues from affecting the homogeneity of group performances.

E. SUMMARY

The present study investigated the effects of test anxiety on discrimination learning. High and low anxiety groups of 30 subjects each were selected on the basis of extreme scores on the Sarason Test Anxiety Questionnaire; groups were then subdivided into matched groups of 15 subjects each. All S's were presented with five of London's graded picture discrimination tests. For half the S's, these tasks were preceded with instructions emphasizing that they were taking a test, while the remaining S's were told that the tasks constituted a game.

Anxiety level and type of directions affected discrimination significantly for three of the tasks presented. In general, both high anxiety and test directions tended to facilitate rapid discrimination. This result was discussed in terms of the apparent alerting affects of anxiety. Where other results obtained, they appeared to be a function of the number of available cues in different types of discrimination tasks, and were discussed in terms of the apparent competing response tendencies of anxious children.

REFERENCES

1. CASTENADA, A., McCANDLES, B., & PALERMO, D. The children's form of the manifest anxiety scale. *Canad. J. Psychol.*, 1955, **9**, 1-6.
2. DOILLARD, J., & MILLER, N. E. *Personality and Psychotherapy*. New York: McGraw-Hill, 1950.
3. FARBER, I. E. The rôle of motivation in verbal learning and performance. *Psychol. Bull.*, 1955, **52**, 311-327.
4. HOCH, P. H., & ZUBIN, J. *Anxiety*. New York: Grune & Stratton, 1950.
5. HULL, C. L. *A Behavior System*. New Haven: Yale Univ. Press, 1952.
6. KELLER, F. S., & SCHOENFELD, W. N. *Principles of Psychology*. New York: Appleton-Century-Crofts, 1950.
7. LONDON, P. Developmental aspects of discrimination in relation to adjustment. *Genet. Psychol. Monog.*, 1958, **57**, 293-336.
8. MOWRER, C. Anxiety and learning theory. In May, R., *The Meaning of Anxiety*. New York: Ronald Press, 1950.
9. OTIS, A. S. *Otis Quick Scoring Mental Abilities Tests*. Yonkers-on-Hudson, N. Y.: World Book.
10. ROSENBAUM, G. Stimulus generalization as a function of clinical anxiety. *J. Abn. & Soc. Psychol.*, 1956, **53**, 281-285.
11. SARASON, S. B., MANDLER, G., & CRAIGHILL, P. G. The effects of differential instructions on anxiety and learning. *J. Abn. & Soc. Psychol.*, 1952, **47**, 561-565.
12. SARASON, S. Measurement and correlation of test anxiety in children. *USPHS Project M-712(02)*.

13. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, **48**, 285-290.
14. ———. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303-320.

c/o London

University of Illinois

1005 W. Nevada

Urbana, Illinois

EXPLORATORY STUDIES OF VERBAL CONTEXT BY MEANS OF CLUSTERING IN FREE RECALL*

Department of Psychology, University of Maryland

RICHARD C. GONZALEZ AND CHARLES N. COFER¹

A. THE PROBLEM

Systematic investigation of verbal behavior and of the learning of connected material must recognize and deal with the important problem of context. It is a matter of common knowledge that the significance of a word, phrase, or even sentence is largely dependent upon the context in which it appears. Despite this, empirical and theoretical inquiry into the rôle of context has traditionally been neglected, although a number of techniques, and studies employing them, have recently appeared which suggest a growing concern with the problem. Miller (12) and, especially, Osgood and Sebeok (15) have highlighted this issue from the point of view of general communication and linguistic theory. Miller and Selfridge (14), working with various approximations to the English language, have shown that the degree of contextual constraint in a passage significantly affects its recall. Deese (7, 8) has indicated that different serial position curves in recall are yielded by variations in contextual organization. King and Cofer (11) have studied the adjective-verb quotient as a factor in learning and retention, and Cofer (3, 4) has used Taylor's "cloze" technique (16) as a means of studying the effects of varying the adjective-verb quotients in prose materials on difficulty of comprehension of the passages. In the more familiar, and perhaps simpler, word association situation, Howes and Osgood (9) and Jenkins and Cofer (10) have found that context can markedly influence the kinds of responses obtained to verbal stimuli. The study of context is thus beginning, by means of a variety of techniques and materials.

The present experiments were undertaken in order to explore in relatively simple situations some limited aspects of context. In relation to the learning and retention of connected materials, we were interested in the kinds of

* Received in the Editorial Office on July 24, 1958.

¹ This Technical Report (No. 19) is based on and summarizes Technical Reports 3, 6, 7, 15, 16, 17, and 18, prepared under a contract between the Office of Naval Research and the University of Maryland, the general project title of which is "Learning, Retention, and Recovery of Meaningful Materials." Reproduction in whole or in part is permitted for any purpose of the United States Government.

effects which adjectival and adverbial modification of nouns and verbs produces. Ideally these effects should be assessed for continuous, connected material, but since the number of simultaneously acting factors in this type of material precludes ready identification and manipulation of variables, it seemed essential to study this problem first in circumstances simpler than that of connected material itself. Information gained in this manner might then allow for more precise investigation of these variables in more complex situations. The situations employed here, however, while much simpler than connected material, were more complex than the usual word association type study.

The series of studies which follow made use of Bousfield's (1) method for studying clustering in free recall. This method involves the initial selection of an equal number of words from two or more mutually exclusive categories, e.g., vegetables, clothes, occupations, etc. The words are arranged in a random sequence to form a single stimulus-word list which is presented to Ss in serial order. The Ss are then asked to recall as many words as they can. When presented such a list Ss tend to recall in sequences of words from the various categories to a greater extent than would be expected by chance. Bousfield has called this phenomenon "clustering" and has developed a technique for its quantification (see section on Method). The successive occurrence in recall of two or more words from the same category is an instance of clustering. This may be illustrated by the following example of a hypothetical recall of a randomized list of words drawn from four categories; Vegetables (V), Animals (A), Occupations (O), and Weapons (W):

V A A A O O W v v v w w w O V o o v v v

The total recall is 20 words, 16 of which occurred in six clusters. Each sequence of words in small caps from a single category represents a cluster.

Since the clustering shown by Ss must involve some addition by them to the randomly ordered stimulus materials, the clustering method has seemed to us, as it has to Bousfield (1), Miller (13) and others, to reflect coding, ordering, mediating, or organizing processes. Such processes, by hypothesis at least, may be involved in learning and retaining connected material and in the influence of context generally. Previous studies using this model have been restricted to the measurement of clustering of single unmodified words. In the studies reported here modifiers were introduced into the situation in order to assess the effects of simple context on retention and organization.

B. METHOD

Features common to the majority of the experiments will be described here.

1. *Subjects*

Ss were male and female college students enrolled in introductory psychology or sociology classes during the regular terms or summer sessions at the University of Maryland or at The George Washington University.² The number of cases in each group is indicated later in the report.

2. *Materials*

Lists of 40 units were used in most of the experiments, as the material to be learned and recalled. The basic control condition consisted of 40 words, drawn 10 each from four categories, the part of speech being ordinarily homogeneous within a list, although varying in different experiments through the categories noun, adjective, verb, and adverb. These words were all common, frequent words, as were the modifiers employed in various experiments. Where modifiers were used, the 40 units were composed of adjective-noun, noun-adjective, adverb-verb, or verb-adverb pairs. In one experiment each unit was composed of four words—adjective-noun-verb-adverb. In all cases the sequence of words was randomized, so as to permit in the random list the number of clusters which Cohen, Sakoda, and Bousfield (6) calculated would occur by chance.

In all instances two recalls were taken, five minutes separating them. This interval was always filled, either by means of a free-association test or by the completion of a "cloze" blank (cf. 3, 16). No words in these interval filling tasks duplicated those in the critical list. In only one experiment was an interfering interpolated task employed (Section G).

3. *Procedure*

Ss were run in groups, frequently in their classes, although in some instances a class was randomly divided and run in different sections. Instructions consisted essentially of the following points (cf. 1): That a list would be read once, following which a written recall would be taken. During the recall the Ss were asked to mark their papers at one-minute intervals on

² Grateful acknowledgment is made to Professors C. E. Tuthill and Eva Johnson, The George Washington University, and to Mr. F. D. Laws and Alice Riddleberger for use of their sociology sections, University of Maryland. We are indebted to Mr. D. J. King, who ran some of our groups and to Carolyn Amundson, Alice Higginbotham, and Marian Wilkins for scoring and tabulating some of the data.

signal from *E*. *E* presented the list of units orally at a regular rate in an even tone without emphasis. In experiments in which units of more than one word were used, *S* ordinarily did not know whether he would be asked to recall the pair (or 4-word unit) or a specified unit member. Five minutes were allowed for this first recall, which was followed by the filled five-minute interval and then by a second five-minute recall, whose conditions were identical to those of the first recall.

4. Scoring

Three of the methods Cohen, Sakoda, and Bousfield describe (6) were used. The first is n , simply the number of words recalled. The second is based on r , or the number of repeated words. The number of repeated words in a cluster is the number of words in the cluster less one, since the first word of a cluster is not a repetition. The r value of the entire recall is the sum of the values for repeated words. In the example cited above (p. 294) r would be 10, since there are 10 repeated words in the six clusters shown (or the total number of clustered words minus the number of clusters).

The second measure is the ratio of repetition (RR). This is essentially the percentage of words recalled that are in clusters and is given by the formula, $RR = \frac{r}{n-1}$. The chance value for RR given by Cohen *et al.* for a 40-item, 4-category list is 0.231 (6, Table 2, p. 4). We have transformed our RR values by means of the arc-sin transformation; the transformed value for the chance RR value is 28.73. This score has been averaged in our groups and is shown in our Tables as $\bar{X} RR$. Also included is a third measure, the percentage of subjects in each group who reached or surpassed the chance value of RR (28.73). This is shown as % RR .

In most of the statistical analyses, t tests, corrected as necessary for repeated measurements, were used to evaluate differences.

C. PRELIMINARY EXPERIMENT

In this first experiment, we wished to explore a methodological point, as well as obtain some preliminary information concerning the effects of adjective modification of nouns or noun clustering. In order to study adjective context, it was, of course, necessary to present adjective-noun pairs. This raised the question as to whether the pairs should be recalled, or only the nouns; as pair recall itself is presumably a more difficult task than noun

recall, we thought it necessary to include both pair recall and noun recall groups in order to evaluate this factor.

Our initial thinking reflected the grammarians' notion that adjectives *limit* the meaning of the nouns they modify and would therefore act to reduce clustering. This required an unmodified list of nouns to serve as a control condition. It seemed also that if an adjective was not appropriate to the noun it modified it might have effects on clustering, different from those it would have as an appropriate modifier; interference effects were thought possible.

To study these several points, five groups of Sociology students were used in the following ways: Group I heard a list of appropriate adjective-noun pairs and recalled the pairs; Group II heard the same pairs but recalled nouns only; Group III heard inappropriate adjective-noun pairs and recalled the pairs; Group IV heard the same pairs but recalled only nouns; Group V, the control, heard and recalled the nouns only. As a part of the initial instructions for this experiment³ *S* was told what he would be asked to recall—pairs or the "second members" of the pairs. The interval between recalls was filled with the free association test.

The 40 nouns used for all groups represented the categories Clothes, Vegetables, Animals, and Occupations. An appropriate adjective was selected for each noun; some effort was made to restrict the selection of each adjective so that it would most appropriately modify nouns in only one category. Tables *A* and *B*⁴ display these pairings and those with inappropriate modification; the latter were formed by shifting the same adjectives to modify other nouns in the list. Scoring in all cases and the measures presented were based on the nouns; the modifiers were ignored.

Table 1 displays a summary of the design of this study and of the chief results. On immediate recall for \bar{X} *RR*, Group III was significantly inferior to all the other groups ($p \leq .05$) and Group I was significantly inferior to Group V. In Delayed Recall, for this measure, Groups II and V did not differ significantly; Group II was significantly superior to Groups I, III, and IV and Group V to Groups I and III; Groups I and IV did not differ

³ This was the only experiment in which *S* knew at the time of list presentation which member of a unit, in this case pairs, he would later be asked to recall.

⁴ Material supplementary to this article has been deposited as Document number 6043 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting \$1.25 for photoprints, or \$1.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

TABLE 1
DESIGN AND RESULTS OF PRELIMINARY EXPERIMENT ON ADJECTIVE MODIFICATION

Group	Modifiers	Stimulus list	What recalled	What scored
<i>Design</i>				
V		Nouns	Nouns	Nouns
II	Appropriate Adj.	Adj-Noun pairs	Nouns	Nouns
I	Appropriate Adj.	Adj-Noun pairs	Pairs	Nouns
IV	Inapprop. Adj.	Adj-Noun pairs	Nouns	Nouns
III	Inapprop. Adj.	Adj-Noun pairs	Pairs	Nouns
<i>Results</i>				
<i>Immediate Recall</i>				
	Group	N	\bar{X} RR	n % RR
A	V	26	40.57	18.04 100
	II	27	38.52	16.96 85.1
B	IV	31	36.44	13.39 87.0
	I	31	31.67	11.58 58.0
C	III	18	21.47	6.22 50.0
<i>Delayed Recall</i>				
			\bar{X} RR	n % RR
A			43.49*	17.65 100
			48.60*	17.33 85.10
B			40.59	13.71 87.00
			34.79	12.33 66.77
C			22.58	5.94 44.44

* Significant increments in \bar{X} RR in delayed recall at .05 level of confidence.

significantly but both were superior to Group III. The results for n and % RR essentially parallel those for \bar{X} RR.

From these differences, together with others to be mentioned in a moment, it appears that, in terms of effects, three groupings have emerged in these findings. Groups V and II are similar (Group A): both show high clustering and recall, and both show a significant increase in amount of clustering with delayed recall. The fact that Group II did not differ significantly from Group V indicates that modification of nouns by appropriate adjectives had no appreciable effect on either clustering or recall. Groups I and IV form another grouping (Group B): both show reduced recall and reduced clustering, as compared to Group A, the differences becoming more apparent in delayed recall (where they are statistically significant); Groups I and IV show no significant increase in clustering with delay, as did Groups II and V. The effects in Group I apparently reflect the influence of the reinstatement of the appropriate modifiers at recall in pair recall but whether this is due to adjective limitation on clustering or to the task of pair recall or some combination cannot be discerned in these data. Group IV, however, is a noun recall group, so that its relatively poor performance derives from the influence of the inappropriate adjectives at the time of list presentation. Hence, adjectives can impair recall and clustering either during list presentation or at recall.

Group III forms the third grouping (C); its poor performance reflects the combined effects of inappropriate adjective modification at presentation and of pair recall. Clustering and recall sink to very low levels here, and the N of 18 for this group reflects only those Ss in the group who produced scoreable records under these conditions; eight papers could not be used. To determine whether these results were artifactual, another Group III was run, consisting of a summer school psychology class, whose members were more highly motivated and at higher college class levels than the original Group III. The 30 Ss of this Group (IIIa) gave a \bar{X} RR for Immediate Recall of 36.35, but this sank to 29.26 at Delayed Recall, a highly significant decline. Evidently these conditions (inappropriate modification combined with pair recall) produce marked disruption of the mechanisms of clustering and recall, especially over a short delay period.⁵

These various groups have demonstrated that adjective modifier effects can be shown on clustering and recall of nouns, although the main effects demonstrated here were accomplished through inappropriate modifiers and through pair recall.

These findings, together with those from one additional group run in this preliminary experiment (see next section), suggested that the kind of modifier effects with which we are dealing are quite complex and that a series of analytic experiments would be necessary to uncover some of the processes in operation. We turn next to consideration of some of the complex processes perhaps involved.

D. MEDIATIONAL PROCESSES IN CLUSTERING

As a part of the preliminary experiment, another group of comparable Ss was run under very different conditions from those described in Section C. It had occurred to us that perhaps adjectives would themselves cluster. The list for this group was therefore composed of 20 nouns and 20 adjectives from the lists in Tables *A* and *B* (Documentary Institute), presented in a randomized serial order rather than as pairs. The first analysis of the recall data was based on clustering in two categories, i.e., nouns and adjectives. The (transformed) values for mean RR under these conditions for immediate and delayed recalls were, respectively, 45.78 and 48.62 (the transformed

⁵ A second Group II was also run, with Ss like those used for Group IIIa. The 29 Ss in Group IIa confirmed the results of the original Group II, yielding mean RR values of 44.78 for Immediate Recall and 48.80 for Delayed Recall, a highly significant increase. While the initial absolute recall and clustering scores of Groups IIa and IIIa were above those for the corresponding original groups, the performance at delayed recall confirms the patterns established on the original groups.

chance value in this case is 44.25). There is thus slight clustering on this two-category part of speech basis. The nouns alone were scored for clustering in terms of the four categories (five nouns in each) they represented, and they yielded mean *RR* transformed values of 33.80 and 35.96 for immediate and delayed recall, respectively, as against the chance value of 27.35. There were then, in this situation, two clustering tendencies: a weak one for parts of speech to cluster and a stronger one for nouns to cluster. However, a rough comparison with the data for the control group in the preliminary experiment was made by doubling the values obtained for the absolute recall of the 20 nouns (*n*) in this experiment. When this was done, the values obtained for immediate and delayed recalls were, respectively, for mean *RR* 33.80 and 35.96, for *n* 16.64 and 15.50, and for % *RR* 75 and 75. The clustering values (mean *RR* and % *RR*) here resemble those for Groups I and IV (Grouping *B*) of the Preliminary Experiment, rather than those of Grouping *A*, although *n* is comparable to that for the latter group. This result suggested to us that some kind of interference effect might be operating and that it might also have operated in the preliminary experiment.

Two further experiments were run in order to explore the possibility that clustering tendencies, associated with either the adjectives or the nouns, might affect the properties of the recall of words in the other category.

In the first experiment the materials used were those shown in Table *A* (*DI*). However the control group in this experiment was presented only the *adjectives* from that list, in the same order, although, of course, without the nouns. There were 25 *SS* in the experimental and 19 *SS* in the control groups. The experimental group was not told at the time of list presentation what it would be asked to recall. It was, in fact, asked to recall the first members (the *adjectives*) of the pairs; the control group, of course, heard and recalled only *adjectives*.

Clustering was scored for the *adjectives* on the basis of the noun categories the *adjectives* modified in the experimental group. The mean *RR* values for the control group in immediate and delayed recall were, respectively, 29.80 and 32.22. While slight, this shows some adjective clustering through the noun categories, even with the nouns absent. The experimental group showed significantly more clustering than the control on immediate recall, yielding a mean *RR* value of 36.52, and its mean *RR* value remained high, at 37.25, in Delayed Recall. Since the only difference between these two groups lay in the fact that, in the experimental group, the *adjectives* had modified the nouns during list presentation, we must attribute the greater

clustering of adjectives in recall by the experimental than by the control group to a sort of backward mediating process; i.e., clustering in the adjectives was apparently mediated by or through the noun categories of the nouns they had modified.

The second experiment run in this connection was the reverse of that just reported. A set of 40 nouns was developed (Table C, *DI*) such that no two or more nouns belonged to the same category (this could be described as a 40-category list). A list of adjectives from four mutually exclusive categories was developed and in one case was presented as appropriate modifiers of the nouns (Table D, *DI*), e.g., pink-sky. In the other case the nouns were presented first, the adjectives second (e.g., sky-pink). This latter case was designed to study the effect of word order on the mediated clustering being investigated.

Group I heard the list of nouns, Group II the adjective-noun pairs, and Group III the noun-adjective pairs. All groups recalled the nouns but did not know at the time of list presentation, in the case of the pair groups, which class of words they were to recall. Clustering of the nouns in all three groups was scored on the basis of the categories of the adjectives which modified the nouns (in either word order). Table 2 shows the results.

TABLE 2
RESULTS OF THE SECOND EXPERIMENT ON MEDIATIONAL PROCESSES

Group	<i>N</i>	Immediate Recall			Delayed Recall		
		\bar{X} RR	<i>n</i>	% RR	\bar{X} RR	<i>n</i>	% RR
I	30	26.61	15.33	43.33	25.20	14.77	36.67
II	22	30.96	11.01	54.55	34.48	11.36	72.73
III	23	29.61	11.00	69.57	31.55	10.96	65.22

The clustering of Groups II and III shows significantly greater values than that of Group I in delayed recall. Group I, whose list was designed to have no categories, showed below chance clustering (28.73) on both recalls. It is clear that in Groups II and III (which did not differ significantly from each other on either recall) clustering of the nouns was mediated through the categories of the adjectives which modified them and that this effect was not dependent on the "normal" word order, i.e., the adjective preceding the noun.

The demonstration of these effects confirms our hypothesis that modification effects of the kind we have studied are complex. It is probable that univocal descriptions of adjectival modification, like that of its limitation function, are too simple, at least in situations such as those used here.

E. SPECIFICITY, CONFLICT, AND FACILITATION EFFECTS OF ADJECTIVE MODIFICATION

The results of the preceding experiments, indicating that clustering of one class of word can be mediated through the categories of another class, suggest that it should be possible to produce a variety of effects through modification or context. To this end three experiments were conducted in order to determine whether conditions could be established in which the limitation or specificity function of adjectives could be shown, whether clustering could be disrupted by the arrangement of conflict between mediating processes in adjectives and in nouns, and whether clustering could be facilitated by arranging congruence between the mediating processes in adjectives and in nouns.

In Experiment V *A*, the Specificity study, a 40-item list of nouns from four categories (leather, metal, textile, wood) was developed (see Table *E*, DI), and an appropriate and specific adjective (Table *E*) was chosen for each noun. Each adjective was chosen so as to have no relevance to the concept expressed in each of the noun categories. In the pair liquor-cabinet, for example, liquor is appropriate to cabinet; cabinet is in the category "wood," but no other nouns in that category (or in the leather, textile, or metal categories) could be appropriately or sensibly modified by liquor or by any conceptually similar word. Thus each adjective is specific to each noun it modifies, and the adjectives did not fall into a number of categories smaller than their total, i.e., 40.

The experimental group heard the list of adjective-noun pairs and recalled only the noun; the control group heard and recalled only the nouns. Scoring was based on the noun categories. Table 3 summarizes the results of this experiment.

TABLE 3
RESULTS OF THE SPECIFICITY STUDY (EXPERIMENT V, A)

Group	N	Immediate Recall			Delayed Recall		
		\bar{X} RR	n	% RR	\bar{X} RR	n	% RR
Control	25	32.07	15.35	73.08	38.85	14.73	96.15
Experimental	26	26.70	9.63	53.33	26.89	9.33	60.00

The control group was significantly superior to the experimental group (which fell below chance level of clustering) at both recalls on all measures. Thus clustering, as well as recall, is significantly impaired by adjective modification, when the adjectives themselves do not cluster and are highly specific (but appropriate) to the nouns they modify. This indicates that

adjectives may limit the meaning of the nouns they modify. Such a limitation presumably hinders or precludes the development of the concepts (wood, metal, etc.) which, in the control group can produce significant clustering and which, also, can apparently facilitate absolute recall for the nouns themselves.

Experiment V *B* was designed to explore the effects of setting the clustering processes in adjectives and in nouns in opposition to each other. A 40-item noun list was prepared from the categories clothes, vegetables, animals, and weapons, and 40 adjectives from color, nationality, space, and deterioration categories were selected (see Table *F*, *DI*). One control group heard and recalled only the noun list, and the other control the adjective list. Two experimental groups were used; both heard the identical adjective-noun pairs, but one recalled nouns only and the other the pairs. In order to set the adjective clustering tendencies in conflict with the noun clustering tendencies for the experimental groups, the adjective-noun pairs were arranged so that the nouns from any one noun category would be modified by adjectives from all of the adjective categories. Thus two or three of the nouns in the clothes group would be modified by color adjectives, two or three by nationality adjectives, and so on with the spatial and deterioration adjectives. Similarly, in each other noun category, adjectives from all the categories would be used to modify the nouns. Hence, the adjective-noun clustering, based on the noun categories, was set in opposition to adjective clustering, based on adjective categories. Again in the case of the presentation of pairs, *S* did not know until time for recall which pair member he would be asked to recall. These results (Table 4) show large amounts of clustering and recall, in both immediate and delayed recall, for the adjective control list (Group 1) and the noun control list (Group 2). By themselves, then, both the adjective and the noun lists clustered to a high degree, the adjective group being superior to the noun group.

Groups 2 and 3, which both represent noun recall, show the effects of the conflicting adjective clustering tendencies. Group 2, which had no adjectives, is superior on both recalls to Group 3 ($p < .05$) on mean *RR*; it recalls significantly more words (*n*) on the first recall than Group 3 ($p < .05$) and the difference is almost significant for the second recall. Clearly, the mediating tendencies in the adjectives have impaired both clustering and recall, even though the adjectives were present only during list presentation. The additional factor of pair recall (Group 4) still further depresses the scores for noun clustering.

The data for Group 4 were also scored for clustering on the basis of the

TABLE 4
RESULTS OF THE MEDIATED CONFLICT AND FACILITATION STUDIES
(EXPERIMENT V B AND C*)

Group	Stimulus	Recall	N	Immediate Recall			Delayed Recall		
				\bar{X} RR	n	% RR	\bar{X} RR	n	% RR
1	adj list	adj	31	46.62	19.63	96.67	52.87	18.87	100
2	noun list	nouns	30	40.60	16.29	87.50	45.94	16.00	87.50
3	adj-noun pairs	nouns	30	34.80	13.26	88.88	29.15	13.81	85.18
4	adj-noun pairs	pairs	24	29.17	10.04	54.16	34.80	10.28	80.00
5	adj-noun pairs	nouns	28	45.66	17.26	96.29	47.96	17.20	92.59
6	adj-noun pairs	pairs	26	36.50	11.72	68.00	44.40	12.04	96.15

* Note Groups 1-4 represent Experiment V B, and the mediating tendencies established in Groups 3 and 4 were in conflict. Groups 5 and 6 represent Experiment V C, and the mediating tendencies here were set in support of each other. Group 2 serves as a control for noun recall and Group 1 for adjective recall in both experiments.

adjective categories. Slightly higher scores were obtained for mean *RR* when this was done than are shown in Table 4 for this group. The new values are 31.74 and 34.90 for Immediate and Delayed Recall, respectively. While not significant, these differences may reflect the greater clustering in the adjective control group (Group 1) as compared with that in the noun control group (Group 2). In any event, the double scoring for paired recall in Group 4 and the demonstration of clustering on the basis of both noun and adjective categories indicates that two organizational tendencies are operative in this situation.

This experiment rather clearly indicates that when two strong clustering tendencies are available but are not congruent, the performance of *S* is impaired as compared with his performance on the basis of either clustering tendency alone. That decrement is a function of the specific relationship between the organizational processes is clear from the next experiment.

Experiment V *C* was designed to investigate another obvious possibility: that two strong clustering tendencies arranged to support one another will facilitate clustering.

To accomplish this, the adjectives and nouns used in Experiment V *B* were rearranged so that the adjectives in one category modified nouns in only one category, e.g., color adjectives modified nouns in the clothing category only (Table G, *DI*). Two additional groups were used (5 and 6). Both were presented the same lists of adjective-noun pairs but experimental Group 5 recalled only the nouns and Group 6 the pairs. Groups 1 and 2 of Experiment V *B* served as controls. The results of Groups 5 and 6 are displayed in Table 4.

Perhaps the most important comparison to be made is that between Groups 5 and 2. Group 5, in which the adjective-noun pairings were set up in a facilitation arrangement is significantly superior to Group 2, the noun control group, in mean *RR* for Immediate Recall and is superior also in all measures in Delayed Recall. Clearly the congruence of the adjective and noun clustering tendencies in Group 5 has facilitated clustering. Group 5 is also highly superior to Group 3, the noun recall conflict group, in all measures at both recalls. Since the same nouns were used for all three groups (2, 3 and 5) and the same adjectives in Groups 3 and 5 (no adjectives were used in Group 2), it is clear that the differences must arise in the effects of the differing arrangements of the adjectives: their conflict with the noun categories depresses clustering in Group 3, as compared to the control, and their congruence with the noun categories increases it, again with reference to the controls.

Similar effects are seen in pair recall, in Groups 4 and 6. Group 4 shows little clustering, because of the combined effect of the conflicting adjectives at both the time of list presentation and at the time of reinstatement of the pairs at recall. Group 6, however, in clustering though not in n , is significantly superior to Group 4 at both recalls, an effect presumably due to the facilitative effect of the congruent adjective-noun pairings. It is also noteworthy that Group 6, the facilitated pair recall group, does not differ significantly on Immediate Recall from Group 3, the conflict noun recall group. This is unusual, because it is the only case in which a pair recall group has equalled the performance of a noun recall group. Even more important is the finding that clustering in Group 6 is significantly superior to that in Group 3 at delayed recall; the facilitation paradigm permits a better clustering performance in the more difficult pair recall task than in the noun recall task, when the latter is subjected to conflict. In addition, the two groups differ but little in n , a measure in which pair recall groups are ordinarily markedly inferior to single word recall groups, presumably because they hear and try to recall twice as many words as do the latter groups.

The three experiments of this section have indicated that under the conditions and in the situations employed, contextual or modifier effects can facilitate the organization of recall shown by clustering and can hamper such organization either by the specificity of their relation to the words modified or through conflict engendered by their own clustering tendencies in relation to those of the words modified.

F. CLUSTERING IN VERBS, ADVERBS, AND FOUR WORD SEQUENCES

The demonstration in the preceding sections of the several effects that adjective modifiers may have on clustering and recall, and especially the facilitation effect described in the immediately preceding section, suggested immediately that exploration should continue in somewhat more complex circumstances. We were especially impressed with the good performance of the facilitated pair recall group, whose absolute recalls (scored in terms of pairs) was as good as that of one noun recall group, which had only half as many words to remember as the prior group. The facilitated pair recall group recalled a considerably greater absolute number of words; we wished to investigate further the effectiveness of the facilitation paradigm in aiding recall. To this end, we wished to study clustering and recall in four word groups, composed of an adjective, a noun, a verb, and an adverb. Before this could be done, of course, it was necessary to study clustering in adverbs,

verbs, and adverb-verb pairs. The first experiment to be reported in this section was concerned with this preliminary question; the second involves the results from the four word sequences.

Our first experiment required that a list of 40 verbs, drawn from four categories, and a list of 40 adverbs, drawn also from four categories, be set up. We found it considerably more difficult to set up specific, mutually exclusive categories for each of these two parts of speech than for adjectives and nouns. Categories of action (verbs) and categories which modify action (adverbs) seem to be more abstract or less denotative than is the case with nouns and adjectives. We report this simply as a part of our experience, recognizing that this point may be due to the special requirements of our situation rather than to a general feature of the language.

The verbs we selected (Table II, *DI*) were categorized in cooking (e.g., baked, fried), activity (e.g., ran, worked), and destroying (cut, blasted) categories and a fourth which represents something that clothes can do (shrink, stretch, etc.), and the past tense was used in all categories. (As we already had nouns and adjectives whose clustering tendencies were known, we selected verbs which, in a facilitation paradigm, would be appropriate to the categories of these nouns. This obviously made the task of obtaining categories for verbs unusually difficult.)

The adverb categories had, of course, to be appropriate to the verb categories selected, as throughout this section we have used the facilitation paradigm. The categories used were temporal (e.g., afterward, lately, soon), speed (e.g., suddenly, slowly, rapidly), quantity (e.g., quite, highly, less), and sound (e.g., mutely, noisily, shrilly). Where the adverbs and verbs were set up in pairs (in either order), the adverbs of the temporal category were paired with the verbs indicating what clothes do, the speed adverbs with the activity verbs, the quantity adverbs with the cooking verbs, and the noise adverbs with the destroying verbs. A pair from each set follows: bulged lately, ran slowly, baked decidedly, exploded mutely. (These would be reversed to make adverb-verb pairs.)

Four groups were used in this experiment: Group I, which heard and recalled verbs only, Group II, which heard and recalled adverbs only, Group III, which heard verb-adverb pairs and recalled adverbs, and Group IV, which heard adverb-verb pairs and recalled verbs. Groups I and II may be referred to as verb control and adverb control groups, respectively; Group III was designed to facilitate recall and clustering of adverbs, and Group IV to facilitate clustering and recall of verbs.

Table 5 displays the results of this experiment. They indicate that the verbs and adverbs (Groups I and II) show significant clustering in both immediate and delayed recall, with the verbs showing higher mean values than the adverbs. These results parallel those for nouns and adjectives. Group III, in which the verbs preceded the adverbs in pairs also shows significant clustering at both recalls but was not significantly superior to its control (Group II) at either recall, although its clustering scores did exceed those of the control. It showed a reverse effect for *n*, as, in both recalls, it scored lower on *n* than its control (these differences were not significant). We have not then demonstrated clear facilitation of clustering in adverbs as an effect of their being "modified" by verbs, and the decline in absolute recall is not characteristic of the facilitation paradigm (see Section E), either.

An even more surprising result appeared with Group IV (adverb-verb, verb recall). The mean *RR* value for Group IV at Immediate Recall is significantly inferior to that for its control (Group I); the value at Delayed Recall continues to be inferior, though not significantly so. At both recalls the difference in *n* (Group IV being inferior) approaches but does not reach significance. Hence, in this comparison, we have not only failed to obtain facilitation but have actually obtained the reverse effect, even though a facilitation design was used.

We remarked above that the verb and adverb categories were not as apparently satisfactory as were the categories for nouns and adjectives. While this might account for our failure to obtain real facilitation in the comparison of Groups II and III, it seems hardly likely that this could, at the same time, account for the decline in performance observed in Group IV. A factor that is present and may account for the difference in results for Groups III and IV, however, is this. In Group III the adverbs (second members of the pairs) were preceded by words (verbs) which in their control condition (Group I) showed more clustering than did the adverbs (Group II). In Group IV, these conditions are reversed, since the adverbs there modify the verbs. It may be that for the facilitation effect to appear, the clustering factor in the first pair member may have to be stronger than it is in the second pair member. This was true in the facilitation study of Section V. We have attempted to check this hypothesis by carrying out a further experiment, one of whose conditions would be adjective-noun pairs in a facilitation design but with the adjectives clustering (alone) less than the nouns (alone). Several lists of adjectives have been studied; to date we have been

TABLE 5
RECALL AND CLUSTERING FOR LISTS OF ADVERBS, VERBS, AND ADVERB-VERB PAIRS

Group	Recall	N	<i>Immediate Recall</i>			<i>Delayed Recall</i>		
			\bar{X} RR	n	% RR	\bar{X} RR	n	% RR
I	verbs	25	38.95	13.67	83.33	40.45	12.46	87.50
II	adverbs	24	35.22	11.76	76.00	36.79	10.68	84.00
III	verb-adv.	24	37.26	9.21	83.33	42.02	8.25	91.67
IV	adv.-verb	25	31.46	7.36	64.00	35.96	7.64	76.00

unable to develop an adjective list which shows less clustering (when it contains categories, as it must in the facilitation design) than a noun list.

We had intended to study clustering and recall of four-word units, in which the adjective-noun and the verb-adverb components each involved the facilitation condition. This experiment was run during the same time period, as, with other *Ss*, the preceding experiment was being carried out. Hence, the four word unit design was run without knowledge of how the facilitation design would actually work in adverbs and verbs. In view of the results of the latter experiment, we might not have expected the kind of facilitation in the four word units that we had originally anticipated, and the results of the four word unit experiment are as a result not entirely clear as to their interpretation. The adverbs and verbs used in the four word units were placed in the order verb-adverb which showed no clear facilitation in the prior experiment, although it showed no loss of clustering, either.

Four word units were made up, each consisting, in order, of an adjective, a noun, a verb, and an adverb. All color adjectives modified nouns of one category (clothing), and with these nouns only the verbs indicative of things clothes do were used; these verbs were followed only by temporal adverbs. In other words the categories of the four parts of speech were set into a facilitation design throughout. The four word units are illustrated by the following instances, one from each of the four sets (see *DI*, Table I for full list): red cloak bulged lately; distant jaguar jumped suddenly; rotting spinach burned much; Russian carbine shot softly. It may be seen that these units do not constitute elegant English.

Group I was presented the list of 40 four word units once, and the list was presented to Group II three times. Groups I and II were asked to recall the four word units. Group III was given the same list once but was asked to recall the nouns only. Group IV was presented once with 20 of the four word units, five from each category, and was asked to recall the units. Scoring was for nouns, only, and only the nouns were scored even when other parts of the unit were recalled. Table 6 exhibits the results.

It is clear that strong clustering was present in all four groups. We can compare these results roughly to those presented for Groups 2, 5, and 6 in Table 4 of the facilitation study (Section *E*). The 40 four word unit group which recalled only nouns (Group III) compares very favorably with the noun control group of the previous study in clustering, the respective mean *RR* values for immediate recall being 40.60 and 39.54. While this indicates no facilitation (as there is in Group 5, Table 4), there is no disorganization arising from the complex four word situation, either. The facilitation

TABLE 6
RESULTS FOR THE RECALLS OF FOUR-WORD UNITS

Group		N	Immediate Recall			Delayed Recall		
			\bar{X} RR	n	% RR	\bar{X} RR	n	% RR
I	40 4-word units (unit recall)	25	33.98	6.16	64.00	34.95	5.92	73.08
II	40 4-word units (3 presentations)	22	35.08	9.45	77.27	41.96	10.27	86.36
III	40 4-word units (noun recall)	24	39.54	10.71	79.17	41.40	10.91	86.96
IV	20 4-word units* (unit recall)	25	25.90	5.88	72.00	24.46	5.63	70.83

* The chance value for RR in the 20-unit case is 21.05, as compared with 28.73 for the 40-unit case.

group which recalled adjective-noun pairs (Group 6, Table 4) yielded a mean *RR* value of 36.50 in immediate recall. The comparable group from the four-word unit study, i.e., the one recalling units (Group I) gave a mean *RR* value of 33.98 on immediate recall, a comparable value. Neither of these four word unit groups showed the increase in clustering at delayed recall that the groups in Table 4 show, but they did not decline either. Clustering, we may conclude was not disturbed in the four word unit situation, but it was not facilitated (similar comments may be made for Groups II and IV, Table 6).

On the other hand, the absolute recall in the four word unit groups is below that for the groups in Table 4 with which we have compared them. Recall for items, then, would appear to be disturbed in the four word unit situation but organization, as indexed by clustering, is not. This result parallels one obtained in a retroactive interference situation to be reported in the next section.

The experiments in this section have shown that while adverbs alone and verbs alone will cluster, their combination does not produce the kind of facilitation observed with adjectives and nouns, although this conclusion must be a highly tentative one, in view of the considerations presented above. The attempt to observe facilitation, through suitable arrangements of the parts of speech, in a four unit case did not succeed, although organization did not deteriorate under these conditions despite reduced absolute recall. We can come to no clear interpretation of the work with four word units, in part because of difficulties with the adverbs and verbs. However, we now doubt, after this trial, that the four word unit situation is a very useful one for study. Inspection of the units as we developed them reveals that they are rather strange creations; the requirement of congruent modification of one category of nouns by adjectives, and the selection of verbs to go appropriately with the nouns and of adverbs to go with these verbs, set up constraints which, we suspect, makes it virtually impossible to set up as many units as we used here and avoid bizarreness in many of them. Further research to explore the "principles" of adjective modification which were apparent in the results of the experiments reported in Section *E* will probably require another situation, perhaps that of specially planned connected material.

G. FURTHER STUDIES OF INAPPROPRIATE MODIFICATION AND AN INVESTIGATION OF RETROACTIVE INTERFERENCE

In the Preliminary Experiment, reported in Section *C*, it was found that the presence of inappropriate modifiers had a detrimental effect on cluster-

ing and recall. This was especially pronounced in the case of pair recall, in which clustering was below chance throughout or declined markedly over the five-minute retention interval to near chance levels (Groups III and IIIa, Section C). Number of words recalled was low under these conditions, also. Inappropriate adjectives present at list presentation seemed to have some effect on noun recall alone; Group IV (Table 1) scored lower for clustering and for n than either the control or the appropriate adjective, noun recall groups on both recalls, and failed to show a rise in clustering between the first and second recalls.

These findings suggested an interference effect on both organization and recall, arising from the inappropriate modifiers alone at the time of learning or from them together with the pair recall task. This interference is probably different from that shown in the mediated conflict study (Section E, Table 4, Groups 3 and 4). In the latter case a list of high clustering adjectives was used, but their categories cut across the noun categories in the nouns they modified. In the inappropriate modification phase of the preliminary experiment, the adjectives used were not set up in categories as such, although they did show slight clustering by themselves when scored for the categories of the nouns they appropriately modified in the Preliminary Experiment. (See the control group of the first mediation study reported in Section D.) The effects of inappropriate adjective modification, then, seemed, on the basis of the preliminary experiment, to be a sort of specificity effect; it is probably not due to adjective clustering and hence must arise from the unusual or inappropriate relation of a specific adjective inappropriately modifying a particular noun.

It is not uncommon that significant information about complex processes is readily obtained by observing the kinds of factors that will break up such processes. Since inappropriate modification seems to break up organization (clustering) and recall, it seemed worth while to explore this aspect further, and an experiment designed to this end is reported in the present section. In the general field of the study of memory, the retroactive interference design has been profitably employed to obtain understanding of the process of forgetting. Since retroaction in relation to the clustering method has not been studied, the second experiment of the section was designed as an initial investigation of this problem; it does not, however, deal with modifier effects.

In the experiment on inappropriate modification, we tested the effect of the number of inappropriate modifiers in the list on clustering and recall. We had anticipated that perhaps there would be some "threshold" value below which inappropriate modifiers would not hinder retention effects and

above which they would. This expectation was based on the belief that clustering tendencies are strong and that it would take a good deal of interference to break them up. (The retroaction experiment reported below had given a basis for this expectation.) In order to give interference the best chance to be displayed, all groups were asked to give pair recalls.

The forty nouns and adjectives used in the Preliminary Experiment were used for all groups; in Group I, however, the pairs were arranged so that in four pairs there was inappropriate modification and in 36 pairs there was appropriate modification. In Group II eight nouns were inappropriately modified, in Group III 12 nouns, and in Group IV 16 nouns were inappropriately modified. The results may be seen in Table 7.

The outstanding feature of the results of this table is the low value of clustering in all groups in Immediate Recall (only Group II exceeds the chance value of 28.73) and of number of words recalled. In Delayed Recall, only Group III shows substantial change in clustering, and in all groups recall remains at low values. There is no evidence that would justify the suggestion of a functional relationship between number of inappropriate modifiers and disruption of clustering or recall; if there is a threshold for the effect of inappropriate modification, these data would suggest that it falls below four for a 40-pair list.

It is probable that the conditions of this experiment were too severe for it to test for all of the effects in which we were interested. The appropriate adjectives were rather specific to their nouns; the task of pair recall under these conditions is very difficult; apparently a few inappropriate modifiers were able, under these conditions, to disrupt what organization these conditions might permit. Further research, especially with the introduction of inappropriate modification in the facilitation paradigm, is needed to clarify further the process of organization through inappropriate modification.

The second experiment followed the conventional retroactive interference paradigm. A 60-item list (1, p. 231), composed of 15 items from each of four categories was presented once to each of two groups, recalled immediately and recalled again after a filled five-minute interval. In the experimental group ($N = 20$), the interval was filled by presenting a similar 40-word list (5, pp. 2-3) and having it recalled once; this list contained no words identical to those in the first list but two categories (animals and names) were common to the two lists and two were not. The control group ($N = 21$) took a 50-item free association test during the interval.

Certain scoring categories additional to those described earlier in this report were used in the present case. These are: categorical intrusions

TABLE 7
RESULTS OF VARYING THE NUMBER OF INAPPROPRIATELY MODIFIED NOUNS

Group	No. of inappropriate modifiers	N	<i>Immediate Recall</i>			<i>Delayed Recall</i>		
			\bar{X} RR	n	% RR	\bar{X} RR	n	% RR
I	4	28	27.67	10.61	60.7	28.88	9.50	50.0
II	8	27	30.24	11.07	55.5	30.53	10.81	55.5
III	12	29	24.96	9.34	41.4	33.17	9.45	79.3
IV	16	27	28.15	8.59	48.1	29.18	9.30	59.3

(words not on the list presented but belonging to the categories of the list), irrelevant intrusions, and interpolated list intrusions (scoreable only for the second recall of the experimental group).

On first or immediate recall, the n for the control group was 21.00 and for the experimental group 20.24, an insignificant difference. At the second or delayed recall the corresponding values were 19.1 and 15.05; the latter represents a significant ($p < .01$) decline in the experimental group and indicates for recall of items the presence of retroactive interference.

The recalls were analyzed for intrusions, but the incidence of intrusions in all categories was very small. The only kind of intrusion giving for either group on either recall a mean greater than one was the relevant intrusion, and the two groups did not differ significantly on this score. The mean number of interpolated list intrusions in the second recall of the experimental group was 0.35.

An analysis was made of recall according to word category of the original list. The control group recalled from 90 to 96 per cent in the second recall of the words in each category it had recalled in the immediate recall test. The experimental group showed a general depression in the delayed recall of recall scores in all four noun categories, giving percentage values from 65 to 78. The categories shared between the original and interpolated lists showed no more effect of interpolation than did the categories which were represented in the original list alone.

While these results show a general depression of recall scores in the experimental as compared to the control group, demonstrating retroactive interference for recall of items, there are no differences in intrusions or in interference in categories common and not common to the original and interpolated lists.

The mean ratio of repetition scores for the control group at immediate recall and at delayed recall, respectively, were 42.69 and 45.58; this rise is not significant. The corresponding scores for the experimental group were 38.17 and 45.32, neither differing significantly from the parallel control group mean, although the rise at second recall for the experimental group is significant at $p = < .01$. It is clear that organization, as reflected in clustering, was not disrupted by the interpolated interference, which, on the other hand, did affect recall for items. This result is similar to that reported in Section *F* for recall and clustering in the four word units. The significance of this correspondence, however, is not immediately apparent.

The experiments of this section have shown further examples of interference in the free recall situation. Minimal amounts of inappropriate modi-

fication disturb clustering and recall when combined with pair recall and highly specific adjective modification. Retroactive interference disturbs recall of unmodified items but not clustering. An obvious further line of inquiry would combine the retroaction design with a variety of modification designs. Such experiments could well result in a significant increase in our understanding of these processes.

H. DISCUSSION

Our impression of the process of recalling the kinds of lists which permit clustering is, on the basis of our work, like this. The subject presumably discovers that the words in the list fall into categories, and his task of recall is facilitated if he used the categories (perhaps their names) as a sort of code which stimulates him to recall appropriate items. As the number of relevant intrusions is usually rather small, the *S* must also be able to discriminate those appropriate words which were in the original list from those which were not. This formulation is certainly incomplete, because on a sheer coding basis one might expect *all* of the words from a category to be emitted in one grand cluster, before the next category was emitted, and so on. As clusters tend to be relatively small (typically two to six words), other factors than coding are probably involved (cf. 2), including a sort of word response strength factor and an inter-word associative factor.

When modifiers enter the picture, they apparently will interact with these coding, response strength, and associative factors. A highly specific adjective may almost "remove" a noun from its category or perhaps change the associative pattern which adheres to the unmodified noun (cf. 10), thus limiting clustering. Perhaps similar effects underlie the disruption arising from inappropriate modification. Such factors may be enhanced in the task of pair recall, which, no doubt, is more difficult than the task of recalling individual words but whose difficulty can be shown to vary as, for example, in the conflict and facilitation designs. It seems to us that in the mediated clustering cases and, also, in the mediated conflict and mediated facilitation instances the results obtained probably reflect on coding operations rather than on associative factors. When the adjective categories cut across the noun categories, but modification is appropriate, there is no *a priori* reason to expect that the adjective-noun association is less strong than it is in the case of congruent adjective and noun categories. This suggests a conflict of coding categories at the time of recall in mediated conflict, and an "over-determination" of the coding situation in the facilitation situation. The clustering of one part of speech (uncategorized) in terms of the categories of

another part of speech with which its words are paired seems to require some kind of mediational process. That uncategorized nouns would tend to appear together because they were modified by a set of color adjectives may be explainable on sheer associative grounds but it is at least reasonable that the adjective category serves as a code permitting the ordering of the discrete nouns. In some situations, we have found clustering to be affected differently from the way in which item recall is affected by certain variables. Thus, in the four word unit situation and in the retroactive interference situation clustering was well preserved, whereas item recall suffered. This suggests that the two processes have some degree of independence.

We began these investigations in the interest of obtaining information that would lead to the identification of variables that could be manipulated in connected material, with whose learning and retention we have primary concern. On the basis of our findings, we suspect that studies of considerable interest can be designed which would utilize the essential features of the specificity, mediated conflict, mediated facilitation, and inappropriate modifier designs in setting up the connected materials to be investigated.

I. SUMMARY

A number of experiments have been reported in which was utilized Bousfield's method for studying clustering in free recall. Bousfield's technique was modified, in many of the experiments, by the use of modifiers for nouns or other parts of speech. The subjects were college students, run ordinarily in groups. A single oral presentation of the list was usually made, followed by an immediate recall, a filled five minute interval, and a second or delayed recall. Some of the chief results and conclusions are as follows:

1. *A specificity effect.* When adjectives are chosen so that they do not belong to mutually exclusive categories and hence do not cluster and are used to modify appropriately only the nouns with which they are individually paired, the result is that the nouns become specific in their meaning and no longer show the clustering which, unmodified, they do.

2. *Mediational effects.* Adjectives and nouns when they belong to a small number of categories show clustering by themselves, i.e., in a list of adjectives alone and in a list of nouns alone. If nouns which do not cluster alone are presented in pairs with adjectives which do cluster, the nouns cluster in terms of the adjective categories. When adjectives which do not cluster alone are paired with nouns which do, the adjectives show clustering when scored for the noun categories. Adverb and verb lists can be developed which independently show clustering.

3. *Facilitation effects.* When clustering adjectives and clustering nouns were set up in pairs so that the adjective categories and noun categories were congruent with each other, clustering and recall of nouns were facilitated. Similar facilitation effects were not demonstrated with adverb and verb combinations or with combinations of four words—adjective-noun-verb-adverb. However, with one exception, in these two cases clustering was not disrupted, and item recall in the four word combinations was disrupted.

4. *A conflict effect.* When categorized adjectives were set to modify nouns so that the adjective categories cut across the noun categories, clustering and recall were adversely affected, even though the modification was appropriate.

5. *Interference effects.* (a) Inappropriate adjective modifiers, presented in the original list, have disruptive effects on clustering and on recall of items; this effect is much pronounced, when the adjective-noun pair, rather than the noun alone, must be recalled. In general, the pair recall task is more difficult than the noun recall task and shows effects, also, from appropriate adjective modification. It is likely that pair recall in some of our experiments reveals the specificity function of adjectives, which, in the case of inappropriate modification, is associated with interference. (b) In a retroactive interference experiment, item recall but not clustering was shown to be affected by interpolated interference.

6. A tentative account is offered of the process of recall in this general situation, involving coding, associative and response strength factors. The results of our various experiments are summarized in terms of this account. It is further suggested that meaningful research with connected material can be pursued involving the variables and "principles" of adjective modification identified in these studies.

REFERENCES

1. BOUSFIELD, W. A. The occurrence of clustering in the recall of randomly arranged associates. *J. Gen. Psychol.*, 1953, **49**, 229-240.
2. BOUSFIELD, W. A., & COHEN, B. H. General review of a program of research on associative clustering. Tech. Rep. #11, Contract NONR 631 (00).
3. COFER, C. N. Cloze procedure in the evaluation of prose stories varying in the adjective-verb quotient. Tech. Rep. #13, Contract NONR 595 (04).
4. ———. A study of the learning of two stories by means of cloze technique. Tech. Rep. #14, Contract NONR 595 (04).
5. COHEN, B. H., & BOUSFIELD, W. A. The effects of a dual-level stimulus-word list on the occurrence of clustering in recall. Tech. Rep. #5, Contract NONR 631 (00).
6. COHEN, B. H., SAKODA, J. M., & BOUSFIELD, W. A. The statistical analysis of the incidence of clustering in recall of randomly arranged associates. Tech. Rep. #10, Contract NONR 631 (00).
7. DEESE, J., & KAUFMAN, R. A. Serial effects in recall of unorganized and sequentially organized verbal material. *J. Exp. Psychol.*, 1957, **54**, 180-187.

8. DEESE, J. Serial organization in the recall of disconnected items. *Psychol. Rep.*, 1957, **3**, 577-582.
9. HOWES, D., & OSGOOD, C. E. On the combination of associative probabilities in linguistic contexts. *Amer. J. Psychol.*, 1954, **67**, 241-258.
10. JENKINS, P. M., & COFER, C. N. An exploratory study of discrete free association to compound verbal stimuli. *Psychol. Rep.*, 1957, **3**, 599-602.
11. KING, D. J., & COFER, C. N. Further observations on the recall of stories varying in the adjective-verb quotient. Tech. Rep. #10, Contract NONR 595 (04).
12. MILLER, G. A. *Language and Communication*. New York: McGraw-Hill, 1951.
13. ———. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81-97.
14. MILLER, G. A., & SELFRIDGE, J. A. Verbal context and the recall of meaningful material. *Amer. J. Psychol.*, 1950, **63**, 176-185.
15. OSGOOD, C. E., & SEBOEK, T. A. Psycholinguistics: A survey of theory and research problems. Supplement, *J. Abn. & Soc. Psychol.*, 1954, **49**, Pp. XI + 203.
16. TAYLOR, W. L. "Cloze procedure": A new tool for measuring readability. *Journ. Quart.*, 1953, **30**, 415-433.

Department of Psychology
Bryn Mawr College
Bryn Mawr, Pa.

Department of Psychology
Graduate School of Arts and Science
New York University
Washington Square, N. Y. 3, N. Y.

THE EFFECTS OF FOOD DEPRIVATION AND RESTRICTED ACTIVITY UPON EXPLORATORY BEHAVIOR OF THE RAT*¹

Department of Psychology, University of Texas

R. M. CARR, J. E. OVERALL, R. K. WHITE, AND W. L. BROWN

A. PROBLEM AND HYPOTHESES

Exploratory behavior of the rat may be defined as the locomotor activity of the *S* in an extended test space which contains no rewards of the usual kind such as food or water. In contrast to experiments involving goal-directed behavior in which *E* designates the "correct" response, investigations of exploratory behavior involve observation of a non-specific response which is recorded in terms of distances traversed.

Recent experimentation by Montgomery (6, 8) and Zimbardo and Montgomery (11) has demonstrated that food or water deprivation: (*a*) significantly reduces the amount of exploratory behavior, with the maximal decrement occurring after 24 hours of food deprivation, and also that (*b*) deprivation does not affect the orderliness of behavior. Montgomery (7) has also investigated the effects of activity deprivation. He found that moderate activity deprivation has no effect upon the amount of exploratory behavior in rats, upon the amount of time spent exploring, or upon the orderliness of exploration. Anderson (2) has found no significant correlations between one measure of general activity and several measures of exploratory behavior. Since Hill (5) has demonstrated that prior restriction of activity increases the general activity level of rats as measured by activity wheel performance, it is concluded that exploratory behavior is independent of the general activity level.

Evidence contradictory to Montgomery's findings has been presented by Thompson (10) who found a *decrement* in exploration when the rat's activity is restricted and an *increase* in exploration under food deprivation. Alderstein and Fehrer (1) found an *increase* of 50 to 75 per cent in exploration under food deprivation and also more consistent exploration by hungry than by sated rats.

* Received in the Editorial Office on July 29, 1958.

¹ This work was conducted at the Radiobiological Laboratory of The University of Texas and the United States Air Force, Austin, Texas, and was supported (in part) with funds provided under Contract AF 41(657)-149 and the School of Aviation Medicine.

The present experiment was designed to determine whether food and/or activity deprivation affect either the amount or the orderliness of exploratory behavior. If exploratory behavior is a primary drive, directed by external stimuli, which undergoes a decrement in strength when another primary drive is present (7); and if there is a permanent decrement in exploration after the animal's first experience in the situation (3); then the present authors would hypothesize (a) that there will be a *reduction* in exploration over time, (b) a *decrement* in exploration due to an internal drive arising from restricted activity, (c) a *decrement* in exploration due to the internal drive aroused by food deprivation, and in view of Montgomery's findings (7, 8), (d) no difference in orderliness of exploration between either deprived and not deprived or "confined" and "not confined" animals.

B. METHOD

1. Subjects

The Ss were 36 experimentally naïve male albino rats of the Sprague-Dawley strain. They were exactly 90 days of age.

2. Apparatus

The apparatus consisted of an enclosed symmetrical Y-maze constructed of 3/4 inch plywood painted white. Each arm was divided into two sections measuring 12 inches in length, five inches in width, and five inches in height and was covered with hardware cloth. The maze was placed in the center of a 6 foot by 8 foot by 7 foot room painted a uniform black and illuminated by a single 40-watt bulb suspended seven feet above the choice point. The *E* observed the animals in the maze from an adjoining room through a one-way mirror.

3. Procedure

Thirty-six rats were randomly divided into two equal groups. One group (deprived) was placed on a 23-hour food deprivation schedule for one week prior to the actual experiment. The other group (not deprived) was maintained *ad libitum* throughout this week and during the ensuing experiment. All Ss had free access to water in their home cages at all times. During the third and fifth days before the start of the experiment, both deprived and non-deprived Ss were placed in confinement cages measuring 6 x 2 x 2 inches with wooden sides and top and a hardware cloth floor for 24 hour sessions. This procedure was used to adapt the Ss to the situation and to ensure that non-deprived Ss would maintain normal eating habits during periods of confinement.

The experiment was designed in such a way that half of the *Ss* were maintained on a deprivation schedule throughout, while the other half were fed *ad libitum* throughout. Whereas each *S* was either "deprived" or "not deprived" on each of the two experimental days, each *S* was "confined" preceding one experimental period and "not confined" preceding the other. The "order" of confinement and non-confinement was counterbalanced so that one-half of the *Ss* were "confined" on Day 1 and "not confined" on Day 2 while the others were "not confined" and then "confined" (see Table 1).

Each *S* was placed in the Y-maze at the choice point and allowed to explore for 10 minutes on each of two consecutive days. For each experimental period, a record was made of the number of 12-inch units explored *per minute* and of the order in which units were explored. The criterion for exploration into a maze unit was a full body-length entry into the unit. Neither food nor water was ever present in the maze.

TABLE 1
SCHEMATIC DIAGRAM OF TESTING PROCEDURE

	<i>Confined</i>	<i>Not confined</i>
Deprived	Group I	Group I
Not deprived	Group II	Group II
	<i>Not confined</i>	<i>Confined</i>
Deprived	Group III	Group III
Not deprived	Group IV	Group IV
	First testing situation	Second testing situation

This experimental design permits a precise test of changes in exploration during the 10-minute sessions, a rather precise measure for the effects of "confinement" upon exploratory behavior, and a less precise test for the effects of deprivation upon exploratory behavior. In addition, possible order effects are counterbalanced and may be tested, and the interactions between deprivation, "confinement," and time in maze can be examined.

C. RESULTS

Table 2 presents a summary of the analysis of variance performed on the data. Twenty-three hours of food deprivation failed to produce a significant difference in exploratory behavior when a two-tailed hypothesis is tested. It should be mentioned, however, that the *directional* hypothesis stating that the extent of exploration of food deprived animals will be equal to, or greater than, that of non-deprived animals must be rejected at the .05

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE

Source	df	Mean square	F
Deprivation	1	124.0	4.03
Order	1	59.0	1.92
Deprivation x Order	1	183.0	5.96*
Error (1)	32	30.8	
Confinement	1	56.0	4.48*
Confinement x Deprivation	1	0.0	0.00
Confinement x Order	1	34.0	2.72
Confinement x Deprivation x Order	1	2.0	0.02
Error (2)	32	12.5	
Time	9	559.4	56.96**
Time x Deprivation	9	7.78	0.79
Time x Deprivation x Order	9	6.56	0.67
Error (3)	288	9.83	
Time x Confinement	9	32.67	7.71**
Time x Confinement x Order	9	34.78	8.20**
Time x Confinement x Deprivation	9	5.22	1.23
Time x Order x Deprivation x Confinement	9	7.78	1.83
Error (4)	288	4.24	

* Significance of $P < .05$.

** Significance of $P < .001$.

level of significance. Figure 1 indicates a consistently higher level of exploratory behavior for *Ss* not previously deprived of food.

Those animals whose activity was restricted by close "confinement" for 23 hours preceding a test period evidenced significantly *reduced* exploratory behavior during the 10 minutes of freedom in the maze ($P < .05$). The interaction between "confinement" and time in the maze was highly significant at the .001 level, reflecting the greater effect of "confinement" upon exploratory behavior during the first three minutes in the maze. Figure 2 presents the number of cul entries during each of 10 minutes for previously "confined" and "not confined" *Ss*.

A significant decrease in exploration as a function of time in the maze was observed. The rate of exploration was initially high, declining rapidly to about 50 per cent of its initial values within 10 minutes. This effect, which was significant beyond the .001 level, can be observed for both the "confined" and "not confined" groups in Figure 2.

The measure of orderliness of exploratory behavior was obtained by tabulating the number of times each *S* entered three unlike arms in succession and the number of times he could have done so, then summing both sets of values for each group and converting them into percentages. These results are summarized in Table 3 for the four groups. There is no significant

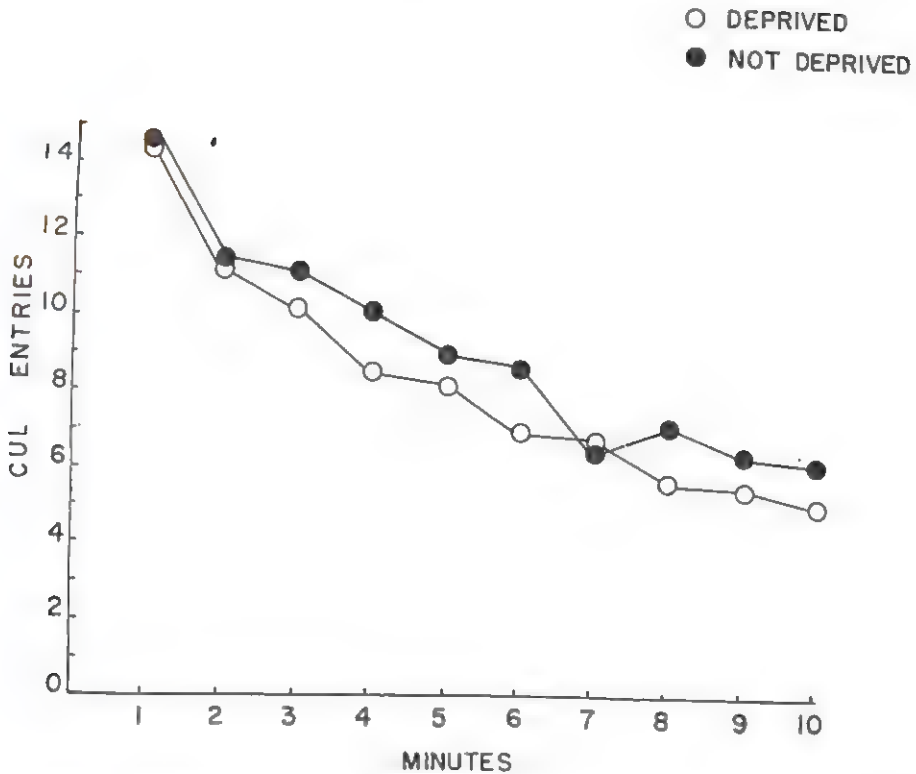


FIGURE 1
MEAN CUL ENTRIES FOR EACH MINUTE OF EXPLORATION FOR DEPRIVED AND NOT DEPRIVED RATS

difference in orderliness of behavior either between the two days of testing or among the four groups.

D. DISCUSSION

The results of this experiment corroborate the hypotheses that (a) there is a *reduction* in exploration over time, (b) there is a decrement in exploration due to restricted activity (confinement), and (c) there is *no difference* in orderliness of exploration between either deprived and not deprived or "confined" and "not confined" animals. The hypothesis of a *decrement* in exploration due to food deprivation is not significantly corroborated when tested by a two-tailed test, but as a directional hypothesis, the effect of deprivation seems to cause a *reduction* in exploration.

The results concerning the effects of deprivation and "confinement" on orderliness of exploration seem to fit in well with those of Montgomery (6,

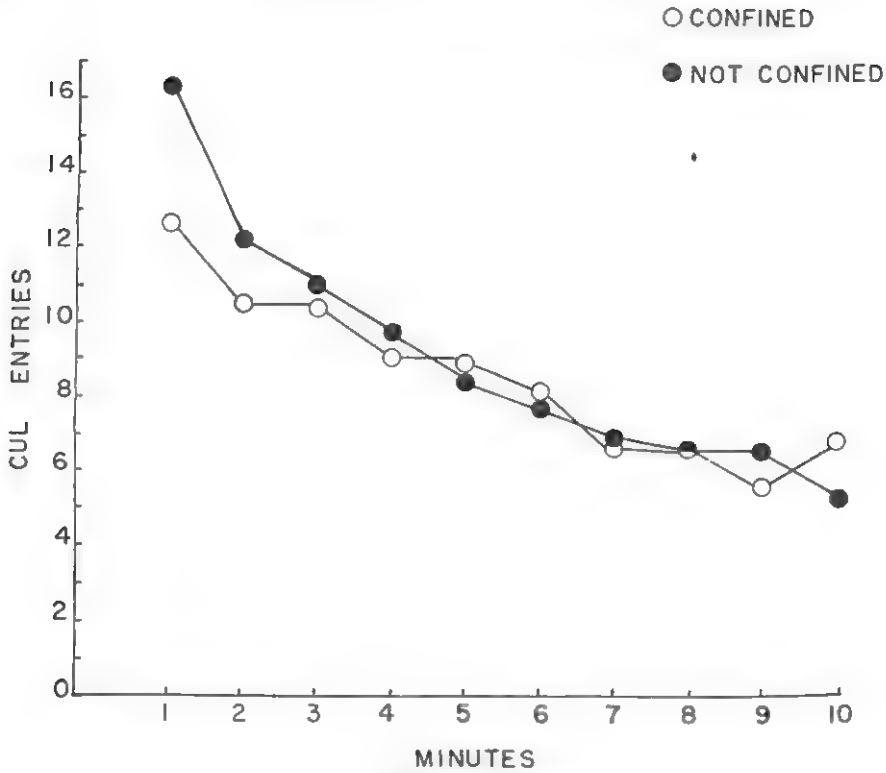


FIGURE 2
MEAN CUL ENTRIES FOR EACH MINUTE OF EXPLORATION FOR "CONFINED" AND "NOT CONFINED" RATS

7, 8), Zimbardo and Montgomery (11), and Berlyne and Slater (4) and, in general, contradict those of Thompson (10) and Alderstein and Fehrer (1). Montgomery (9) has stated that exploratory behavior results from a drive which is highly dependent upon stimuli of the external environment. The present authors believe that the presence of an internal drive aroused by hunger or restricted activity acts as an inhibitor of the exploratory drive. Although the amount of exploration is reduced when internal drives are

TABLE 3
ORDERLINESS OF EXPLORATION

Group	Mean per cent	
	Day 1	Day 2
I	57%	58%
II	55%	55%
III	54%	53%
IV	56%	52%

present, the pattern or orderliness of behavior remains the same so long as the external stimuli which serve to direct the exploratory drive remain the same for both deprived and not deprived and/or "confined" and "not confined" animals. Therefore, the increase in exploration due to food deprivation obtained by Alderstein and Fehrer (1) appears due to the introduction of extra-maze cues rather than due to the deprivation training (11).

The decline in exploration over time lends support to Berlyne's hypothesis (3) in that the rat becomes rapidly satiated by the novel assortment of stimuli present in the maze runways in the first minutes of exploration and in that the original 10-minute exposure is not long enough to cause a permanent satiation and there is a rearousal of novel stimulation in the second, as well as future, presentations of the same maze situation at 24-hour intervals.

E. SUMMARY

Thirty-six experimentally naïve male albino rats exactly 90 days of age were randomly divided into two equal groups. Half of the *Ss* were maintained on a deprivation schedule throughout the experiment. Each *S* was "confined" preceding one experimental period and "not confined" preceding the other. The order of "confinement" and "non-confinement" was counter-balanced so that one-half of the *Ss* were "confined" on Day 1 and "not confined" on Day 2 while the others were "not confined" and then "confined." Each *S* was placed in the Y-maze at the choice point and allowed to explore for 10 minutes on each of two consecutive days. Records were made of the number of 12-inch units explored per minute and of the orderliness of exploration.

The experimental results support the hypotheses that: (a) there is *reduction* in exploration over time ($P < .001$); (b) a *decrement* in exploration due to an internal drive arising from restricted activity ($P < .05$); (c) *no difference* in orderliness of exploration between either deprived and not deprived or "confined" and "not confined" animals; and (d) a suggestion of a decrement in exploration due to the internal drive aroused by food deprivation.

REFERENCES

1. ALDERSTEIN, A., & FEHRER, E. The effect of food deprivation on exploratory behavior in a complex maze. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 250-253.
2. ANDERSON, E. E. The interrelationship of drives in the male albino rat: II. Intercorrelations between 47 measures of drives and of learning. *Comp. Psychol. Monog.*, 1938, **14**, No. 6.
3. BERLYNE, D. E. The arousal of perceptual curiosity in the rat. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 238-246.

4. BERLYNE, D. E., & SLATER, J. Perceptual curiosity exploratory behavior and maze learning. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 228-231.
5. HILL, W. Activity as an autonomous drive. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 15-19.
6. MONTGOMERY, K. C. The relation between exploratory behavior and spontaneous alternation in the white rat. *J. Comp. & Physiol. Psychol.*, 1952, **44**, 582-589.
7. ———. The effect of activity deprivation upon exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 438-441.
8. ———. The effect of hunger and thirst drives upon exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 315-319.
9. MONTGOMERY, K. C., & MONKMAN, J. A. The relation between fear and exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 132-136.
10. THOMPSON, W. R. Exploratory behavior as a function of hunger in "bright" and "dull" rats. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 232-236.
11. ZIMBARDO, P. G., & MONTGOMERY, K. C. The relative strengths of consummatory responses in hunger, thirst, and exploratory drive. *J. Comp. & Physiol. Psychol.*, 1956, **50**, 504-508.

Department of Psychology
University of Texas
Austin, Texas

BOOKS

Now that there is a special *APA* journal completely devoted to the publication of book reviews, it is no longer necessary that other journals emphasize such publication. It has always been our conviction that book reviews are a secondary order of publication unless they carry information that is equally important as the book. However, the publication of book titles is a very important service, and we shall continue to render that service.

In any given issue of this journal, we may continue to publish one or more book reviews, but we do not consider such publication a major function of this journal. In line with this policy, we can no longer pay for such manuscripts.

(Sargant, W. *Battle for the Mind: A Physiology of Conversion and Brain-washing*. New York: Doubleday, 1957. Pp. 263.)

REVIEWED BY GLADYS C. SCHWESINGER

Patterns of individual belief and behavior are built up painstakingly over the developmental period and on into maturity, and from these future behavior can ordinarily be predicted. That is, unless the individual is subjected to such severe stress that his originally existing patterns are dispersed, and new patterns are imposed in their place. Throughout the ages, severe and prolonged stress has been applied to individuals and groups, from ancient to modern times—in political, priestly, and police settings, and more recently in psychiatric medicine. Results are inevitably effective.

What are the techniques for achieving a complete change of behavior and belief whether for good or for evil? They lie in individual reaction to conflict and stress, kept up so constantly and continuously that the brain and nervous system collapses and inhibition sets in. Following breakdown, the subject is in a state of such increased susceptibility that he will subscribe to almost anything, even though it might mean his own death warrant. The revivalist exhortations and techniques of a John Wesley or a Jonathan Edwards, or of the many religious leaders among primitive tribes and voodoo cults; the methods—improperly labeled by laymen as “third degree”—in the hands of the police trying to wring out confessions of guilt; the “brain-washings” of political agents as practiced today in Communist China and Russia; and—like it or not—the psychoanalytic procedures, initiated by Freud and used by many psychiatrists, all have elements in common for changing the mental make-up of individuals being treated.

Furthermore, the experience of all these workers is in direct agreement with the findings of mechanistic research as conducted by Pavlov, which provides convincing physiological explanations and does away with any need for psychological theory, vague or definite, to account for changed behavior. Only now are we becoming aware of the parallels between these various agencies, and it is Sargant who is pointing them up in no uncertain way.

What is the basic structure upon which all these "rehabilitators" depend? What are the elements influencing the effectiveness of this application in individual cases? Briefly they may be summed up as follows, when viewed through Pavlovian experimentation:

(a) Individuals, whether animals or humans, respond to stress and conflict. (b) What is more, they respond in typical ways according to their inherited basic temperament patterns (which incidentally are seen to be in fairly close agreement with the four categories named by Hippocrates, over two thousand years ago) and allowing for gradations in between. Variations in behavior are drawn into detailed expression by environmental agents and influences. (c) When stresses and conflicts are too great for the nervous system to handle, breakdown follows. (d) On complete breakdown, individual normal behavior, as governed by heredity and conditioning, begins to change. (e) Amount of stress varies with physical condition, resistance being lowered by fatigue, fevers, drugs, and glandular changes. (f) "Transmarginal resistance" (i.e., beyond the nervous system's capacity to respond normally) brings about inhibition. (g) Transmarginal inhibition becomes protective, and results in altered behavior, of which Pavlov noted three kinds: the "equivalent" phase in which reaction to both strong and weak stimuli is the same; the "paradoxical" phase, in which greater response is made to weak stimuli than to strong; the "ultraparadoxical" phase in which conditioned responses turn from positive to negative and vice versa. (h) Transmarginal "protective" inhibition in dogs features a state resembling hysteria in man.

It is interesting to note that Pavlov himself, although not a clinician, wrote "neurologists and psychiatrists in their respective domains will inevitably have to reckon with the following fundamental patho-physiological fact: the complete isolation of functionally pathological (at the aetiological moment) points of the cortex, the pathological inertness of the excitatory process, and the ultraparadoxical phase."

The extent to which Pavlov was right in his prediction can be followed in Sargant's comparison of the behavior of animals with that of human beings, especially those many shell-shocked victims of the Normandy beachhead and the London blitzes during World War II. "Protective inhibition" embodied constant fatigue, and repeatedly the doctors noted patients' inability to dis-

criminate important from unimportant stimuli. Excitatory symptoms could be induced on very slight provocation. Emotional exhaustion followed a typical pattern as to time of onset and symptomology. And on down the line, point by point, the psychiatrists, treating war-shocked service-men and civilians, identified states of transmarginal inhibition, as described by Pavlov in his dogs, in its several phases, and the correlative need to temper sedation to temperamental type if relief was to be effective and treatment not actually injurious in some instances.

Psychotherapists use drugs to help release inhibited emotions of terror, anger, frustration and despair, and produce "abreaction"—i.e., the revival of memories and their related emotions—to relieve the personality of their influences. Laymen use alcohol to discharge their repressed feelings in socially acceptable ways. Vigorous dancing, with wild Negro jazz, and drumming and rhythmic dancing, have also been employed by priests and demagogues to produce catharsis. Free association of ideas, dreams, and so on, all lead to the emotional storm of "transfer," and end in the patient's acceptance of the therapist's interpretations of his symptoms.

A variety of shock treatments and leucotomy, concurring with states of excitement in the brain, go on to protective inhibition and collapse, before "cure" is achieved. Even some of the religionists, like John Wesley, were not unaware of the effective use of electric shock in bringing about conversion by similar steps.

Fear of everlasting hell and damnation, presented as threats to those who rejected the evangelist's offer of eternal salvation, aroused to highest pitches of indignation and anger, induced a disturbance of brain function, and brought on a heightened suggestibility in the convert. Or, as the author says they seemed to wipe off the "cortical slate" and reverse earlier behavior patterns.

Among primitive peoples, local and elaborate initiation rites—as described in Frazier's *Golden Bough*—disclose similar use of stress, more stress, seclusion, pain, excitement, fear, uncertainty as to outcome, drums beating, music, rhythm, a play on the emotions to the point of emotional exhaustion, an offer of relief on admission of unworthiness, guilt, confession, etc., all pointing up the same thing, whether to relieve depression and melancholia, to bring about religious conversion, or to induce confessions of guilt (even when not guilty). All are hastened by the agent if he can learn the state of the subject's mind, of what he is thinking, his "sore spots," and then set to work on these.

The stress factors the agent employs usually include: fatigue, disturbances

of sleep over a long period, the creation of feelings of anxiety and guilt, mental conflict, prohibition against the victim's telling anybody else that he is being "investigated," social degradation, the use of stool pigeons who are under instructions to sympathize with the victim, to identify themselves with his problems, to persuade him to confess, and then for the agent to confront the victim with a real or a pretended confession supposedly obtained from an associate who was charged with the same crime. A vicious side-effect sometimes emerges in that the victim actually begins to feel even an affection for the examiner (paradoxical and ultraparadoxical phases approaching!).

To be noted also, in passing, is the observation that the readiest victims of the brain-washing, or the religious conversion, are the simple, healthy extroverts. The hardest—and even the completely impossible—to "cure" are the psychotic, who have no great feelings or need to conform to pressure in the first place!

Research is needed to discover the different stress methods which will prove most effective for different temperamental types. From this point, it looks as if the extrovert could be reached by non-specific group-excitatory methods and strong and repeated emotional arousal, while the obsessional person seems to require physical debilitation, individual approach, and pressure. The psychopaths, many of whom are criminals, with immature brains, appear to need to wait the discovery of some drug which would speed their delayed brain maturation.

How far group-excitatory methods can be applied to everyone, and differential reaction to such things as solitary confinement, the variation in personality types with increase in civilization, (chronically anxious, obsessional, hysterical, schizoid, and depressive "normals") call for variation in group and individual therapies. Research is needed on: the suitability of different fear or anger-provoking psychological stimuli as applied to different types, different environments, etc., and cultures, etc., the finding of the right "sore spots" and their variation in application with educational level and previous conditioning of the persons investigated or worked upon (a favorite device of Noah Porter!).

A crucial practical question, needing answers from the scientist is: How can one avoid or resist political conversion? It is obvious that the powerful physiological and mechanistic techniques used by agents cannot be offset by an individual's will power or even his intellectual awareness of what is happening to him, and that other reactions must be sought. In passing, it should be remembered that the dogs who outwitted Pavlov were those who refused to coöperate with him; correlatively human beings who evade action may not

get brain-washed. If one does not talk in court, gives no sign of damaging confessions, does not represent himself but rather has a lawyer speak for him, and maintains a detached attitude throughout the trial or torture, he is in a better position to resist breakdown, collapse, brain inhibition, and "conversion." If he can muster some amusement and humor over the methods being used against him, it will serve as a good balance against the production of emotional excess within himself.

These are good points to remember if one should ever happen to get caught by police, become involved in a political "purge," or even find himself cornered at a revivalist meeting, but specific, well-directed research in the hands of trained research scientists, using these leads as points of departure, would be much more welcome!

Meanwhile all readers of Sargant's revealing treatise, all citizens of democracies everywhere, and all pretenders to a sound philosophy of education, will readily subscribe to the concluding paragraph of his remarkable contribution to modern knowledge and analytical understanding when he says:

The brain should not be abused by having forced upon it any religious or political mystique that stunts the reason, or any form of crude rationalism that stunts the religious sense (p. 239).

California Youth Authority
P.O. Box 170
Ventura, California

(Anastasi, A. *Differential Psychology*, 3rd Ed. New York: Macmillan, 1958. Pp. 664.)

REVIEWED BY GLADYS C. SCHWESINGER

The first edition of *Differential Psychology* by Anne Anastasi came out in 1937, in two parts, 20 chapters, and on 615 pages. Twelve years later, in 1949, with Foley as collaborator, a second edition appeared in three parts, with 24 chapters and on 894 pages. Now in 1958, with the original author again the sole worker, the book re-appears in one part, in 18 chapters and on 664 pages.

What have been the changes between the last two editions: the additions, the drop-outs, and the shifts in emphases over the last 11 years? Allowing

for the re-inclusion of old topics appearing under new titles, and for regroupings of old material with other than their original companion subjects, the major addition seems to be two-fold, in that an extra chapter has been included on "Heredity and Environment" (Basic Concepts and Methodology); and an extra one on "Trait Organization" (Theories, Methods, and Major Results). The drop-outs appear to include the chapters having to do with (a) "Psychological Testing" (Basic Concepts, Nature and Extent of); and (b) "Factors in Simple Behavior Development" (Biological and Psychological) with the chapter on "Twins and Foster Children" finding a suitable depositary in the discussion of "Family Resemblance." The discussion of "Racial Differences" has been reduced from three chapters in 1949 to two in 1958, while "Bodily Conditions and Behavior" seems to have disappeared altogether.

This re-alignment and reduction brings more compactness and solidity to the latest volume, and leaves to other authors the job of presenting the ancillary material which swelled Anastasi and Foley's second edition. Within the compass of a single book, particularly when dealing with a widely spread and complicated subject, and realizing its dependence on many closely related topics, authors are ordinarily faced with a dilemma: whether to overweigh their book with both, or to cut to one. Usually they cannot cover both, especially if they wish to emerge with clear-cut principles and statements of fact. But authors, like wanderers, can get lost in the forest by being too attentive to the separate trees and it sometimes takes two or three editions to find this out.

Anastasi's third edition carries the mark of material that has ripened enough for it to be better digested by psychologists, perhaps with the added years, and perhaps also with the constantly increasing insight which the experts in this area are acquiring who have labored over it for a long time. When a science is young, or when a particular area of research in it is relatively new, the exponent feels obligated to present supporting evidence for almost every statement he makes. But as time goes by and old evidence gets assimilated, leaving only the results and conclusions to be retained and remembered by readers, a more dogmatic and less evidential type of presentation is in order, and this makes for clearer presentation and retention.

It may be that Anastasi has profited by the gradual growing-up of her subject matter, for here and there, a reader, fingering through the 1958 edition of "Differential Psychology," can find as many as a dozen consecutive pages of discussion which are not adorned by bibliographic references! The ground work has been laid, the plowing and the sowing done in years gone by, and now comes concern chiefly with the harvest! Yet Anastasi is

too good a research reporter and interpreter to lose her authorities completely, and the reader can still count on each chapter being well supplied with extensive relevant and carefully selected bibliographic references. Pictures, charts, graphs, and even photographs still round out most of her discussion.

As a compendium of research findings, on individual and group differences and their underlying basic concepts of study, which any one who is approaching these subjects would need to grasp, the book encompasses about everything the average teacher, student, or research worker would want who has to organize a lecture, make a speech, write a paper, or start a piece of research in this area. In another decade, the story will have to be retold again, for today's and tomorrow's researchers will dig deeper and come up with more pay dirt, perhaps a new kind, to add to our enlightenment and our confusion. Maybe, then, Anastasi will scrap much of what is left of her today's logical arrangements and categories of material and develop a new cross-section approach to the whole thing. By then, say in a dozen years, Edition 4 will not even show its family relationship to Editions 1, 2, and 3. And what emerges in 1970 may be almost as definite, as free from controversy and confusion, and as precise as a textbook. But that time is not yet.

Today and tomorrow we will still have to work through the old approach to obtain a better understanding of individual differences. We know they exist, and in many instances, we can even tell how much, but we cannot yet answer "why" they exist, or "why" they vary, even when certain fundamental conditions are held constant. We are still laboring to find better techniques for studying what we observe and, with improvement in our methods, we will continue to try to find answers that will be more precise.

As for group differences, like individual differences, they are recognized, are studied as to category and degree of existence, but are still loaded with qualifications as to causation and origin. We are still almost blindly grappling with the things that seem to bring about group differences. We scrutinize our tools and our examining instruments and we devise new ones. When by applying these we find ourselves emerging with a wealth of variation that seems a little out of line in this ever so democratic land of "equality of opportunity," we realize that we are still unable to pierce the fog that blankets the impenetrable.

"Why"? Not a scientific question, we were told first by Newton and we heard it often during our junior years of research training, but yet isn't that just what workers in differential psychology are trying to find out? Isn't that the indissoluble residue of today's outlook on psychological differences, whether among individuals or groups?

Anastasi devotes 300 pages to human differences that could be largely ascribed to biology (physique and constitution): to environment (training, age, etc.), and to an almost unravelable interlocking of both (family resemblances), all with qualifications, of course. She gives 50 pages to statistical considerations; some 200 pages to groups which deviate from the norm (mental defectives, genius), or from each other (sex, social class, and race); and then she winds up her huge survey by considering the cultural "frame of reference" and its ramifications as they impinge on the individual and the group, all to achieve the aim of *understanding behavior*, or to answer the question of why people react as they do. It is these ramifications of culture, she concludes, which seem to determine psychological make-up in a way that original nature, biology, or heredity never do. Multiplicity of culture factors should, therefore, be more productive of variation than simplicity. If this is so, then western man should prove more variable than primitive man in his circumscribed cultural medium. I wonder if he is? Or do we know enough about the primitive man as a "range of individuals" even to take a broad guess?

Certainly Pavlov in his later work of mechanistic conditioning of dogs had to face the reality of their inherent native differences in temperament—at least as he recognized four broad basic categories that in essence were somewhat similar to those Hippocrates advanced two thousand years ago. And psychiatrists treating soldiers and civilians for battle neuroses in England and on the continent learned from Pavlov the wisdom of tempering their sedations to the individual native temperament they were dealing with!

A close study of Sargant's findings—presented in a book (that is not included in Anastasi's review) indicates that basic behavior patterns in man—as well as in dogs—are more dependent on inherited higher nervous systems than we are prepared to admit. Sargant presents compelling evidence that temperament, like potential for intelligence, comes wrapped up in the original package, and unfolds only after the content of the package is exposed and put into an inter-relationship with whatever environment it finds itself in. This can lead observant psychologists to sway their balance against Anastasi's position and to be on the alert for further research, as suggested by Sargant, more specific conclusions, and still another edition of *Differential Psychology* a few years hence.

California Youth Authority
P.O. Box 170
Ventura, California

BOOKS RECENTLY PUBLISHED

- ABRAMSON, H. A., *Ed.* *Neuropharmacology*. New York: Josiah Macy, Jr., 1959. Pp. 285.
- ACKERKNECHT, E. H. *A Short History of Psychiatry*. New York: Hafner, 1959. Pp. 98.
- ALBEE, G. W. *Mental Health Manpower Trends*. New York: Basic Books, 1959. Pp. 361.
- AMES, L. B., METRAUX, R. W., & WALKER, R. N. *Adolescent Rorschach Responses*. New York: Harper, 1959. Pp. 313.
- BASS, B. M., & BERG, I. A., *Eds.* *Objective Approaches to Personality Assessment*. Princeton: Van Nostrand, 1959. Pp. 233.
- BLANTON, S. *Now or Never*. Englewood Cliffs, N. J.: Prentice-Hall, 1959. Pp. 273.
- BOSSELMAN, B. C. *Self-Destruction: A Study of the Suicidal Impulse*. Springfield, Ill.: Thomas, 1958. Pp. 94.
- BRAHMACHER, S. *Social and Moral Attitudes*. Bombay: Orient Longmans, 1959. Pp. 113.
- BRIDGMAN, P. W. *The Way Things Are*. Cambridge: Harvard Univ. Press, 1959. Pp. 333.
- CHANCE, E. *Families in Treatment*. New York: Basic Books, 1959. Pp. 234.
- CHAPANIS, A. *Research Techniques in Human Engineering*. Baltimore: Johns Hopkins Press, 1959. Pp. 316.
- COHEN, J., ROBSON, R. A. H., & BATES, A. *Parental Authority: The Community and the Law*. New Brunswick: Rutgers Univ. Press, 1958. Pp. 301.
- COLE, L. *Psychology of Adolescence*. New York: Rinehart, 1959. Pp. 731.
- COMBS, A. W., & SNYGG, D. *Individual Behavior*. New York: Harper, 1959. Pp. 522.
- CYPREANSEN, L., WILEY, J. H., & LAASE, L. T. *Speech Development Improvement, and Correction*. New York: Ronald Press, 1959. Pp. 353.
- DAVIS, RABBI D. L. *Understanding Judaism*. New York: Philosophical Library, 1958. Pp. 119.
- DEESE, J. *The Psychology of Learning*. New York: McGraw-Hill, 1958. Pp. 367.
- DIMOCK, M. E. *Administrative Vitality*. New York: Harper, 1959. Pp. 298.
- ECKARDT, A. R. *The Surge of Piety in America*. New York: Association Press, 1958. Pp. 192.
- EISENSON, J., *Ed.* *Stuttering: A Symposium*. New York: Harper, 1958. Pp. 402.
- EISSLER, R. S., FREUD, A., HARTMANN, H., & KRIS, M., *Eds.* *The Psychoanalytic Study of the Child*. New York: International Univ. Press, 1958. Pp. 573.
- FLEMING, C. M. *Teaching: A Psychological Analysis*. New York: Wiley, 1958. Pp. 291.
- FLINT, B. M. *The Security of Infants*. Toronto: Univ. Toronto Press, 1959. Pp. 134.
- FRAIBERG, S. H. *The Magic Years*. New York: Scribners, 1959. Pp. 305.
- FRENCH, J. L. *Educating the Gifted*. New York: Henry Holt, 1959. Pp. 543.
- FRENCH, T. M. *The Integration of Behavior*. Chicago: Univ. Chicago Press, 1958. Pp. 484.
- GALDSTON, I., *Ed.* *The Family in Contemporary Society*. New York: International Univ. Press, 1958. Pp. 147.
- GALLAGHER, J. R., & HARRIS, H. I. *Emotional Problems of Adolescents*. New York: Oxford Univ. Press, 1958. Pp. 174.
- GARRISON, K. C. *Growth and Development*. New York: Longmans, Green, 1959. Pp. 559.
- GARRISON, K. C., & FORCE, D. G., JR. *The Psychology of Exceptional Children*. New York: Ronald Press, 1959. Pp. 586.
- GOODENOUGH, F. L., & TYLER, L. E. *Developmental Psychology*. New York: Appleton-Century-Crofts, 1959. Pp. 552.
- GRANT, N., JR. *Art and the Delinquent*. New York: Exposition Press, 1958. Pp. 36.
- GRONLUND, N. E. *Sociometry in the Classroom*. New York: Harper, 1959. Pp. 340.
- GROSS, L. *God and Freud*. New York: McKay, 1959. Pp. 215.
- HARING, A., & YODER, W. O., *Eds.* *Trading Stamp Practice*. Bloomington: Indiana Univ., 1958. Pp. 390.

- HARLOW, H. F., & WOOLSEY, C. N., *Eds.* Biological and Biochemical Bases of Behavior. Madison: Univ. of Wisconsin Press, 1958. Pp. 476.
- HARROWER, M. Personality Change and Development: As Measured by the Projective Techniques. New York: Grune & Stratton, 1958. Pp. 383.
- HARSH, C. M., & SCHRICKEL, H. G. Personality. New York: Ronald Press, 1959. Pp. 536.
- HESS, R. D., & HANDEL, G. Family Worlds. Chicago: Univ. of Chicago Press, 1959. Pp. 305.
- HILL, R., STYCOS, J. M., & BACK, K. W. The Family and Population Control. Chapel Hill: Univ. of North Carolina Press, 1959. Pp. 431.
- HIMMELWEIT, H. T., OPPENHEIM, A. N., & VINCE, P. Television and the Child. New York: Oxford Univ. Press, 1958. Pp. 522.
- HONIGSMANN, J. J. The World of Man. New York: Harper, 1959. Pp. 971.
- HOOK, S., *Ed.* Psychoanalysis, Scientific Method, and Philosophy. New York: New York Univ. Press, 1959. Pp. 370.
- HOWE, R. L. The Creative Years. Greenwich, Conn.: Seabury Press, 1959. Pp. 239.
- HURLOCK, E. B. Developmental Psychology. New York: McGraw-Hill, 1959. Pp. 645.
- HUTT, M. L., & GIBBY, R. G. The Child: Development and Adjustment. Boston: Allyn & Bacon, 1959. Pp. 401.
- JACO, E. G., *Ed.* Patients, Physicians, and Illness. Glencoe: Free Press, 1958. Pp. 156.
- KARLSSON, G. Social Mechanisms. Glencoe: Free Press, 1958. Pp. 156.
- KARPMAN, B. Child and Juvenile Delinquency. Washington, D. C.: Psychodynamics Monograph Series, 1959. Pp. 364.
- KATO, H. Japanese Popular Culture: Studies in Mass Communication and Cultural Change. Rutland: Charles E. Tuttle, 1959. Pp. 223.
- KIELL, N. The Adolescent Through Fiction. New York: International Univ. Press, 1959. Pp. 345.
- KIRK, S. A. Early Education of the Mentally Retarded: An Experimental Study. Urbana: Univ. of Illinois Press, 1958. Pp. 216.
- KOCH, S., *Ed.* Psychology: A Study of a Science. New York: McGraw-Hill, 1959. Pp. 710.
- LAFITTE, P. Social Structure and Personality in the Factory. New York: Macmillan, 1958. Pp. 228.
- LANE, R. E. Political Life: Why People Get Involved in Politics. Glencoe: Free Press, 1959. Pp. 374.
- LEVY, D. M. Behavioral Analysis. Springfield: Charles C Thomas, 1958. Pp. 370.
- LIPSET, S. M., & BENDIX, R. Social Mobility in Industrial Society. Berkeley: Univ. California Press, 1959. Pp. 309.
- LIVERIGHT, A. A. Strategies of Leadership. New York: Harper, 1959. Pp. 140.
- LOREE, M. R., *Ed.* Educational Psychology. New York: Ronald Press, 1959. Pp. 425.
- MAGOUN, F. A. Successfully Finding Yourself and Your Job. New York: Harper, 1959. Pp. 238.
- MALCOLM, N. Dreaming. New York: Humanities Press, 1959. Pp. 128.
- MANDLER, G., & KESSEN, W. The Language of Psychology. New York: Wiley, 1959. Pp. 301.
- MASLAND, R. L., SARASON, S. B., & GLADWIN, T. Mental Subnormality. New York: Basic Books, 1959. Pp. 442.
- MARTIN, J. B. The Pane of Glass. New York: Harper, 1959. Pp. 397.
- MCCARTHY, R. G., *Ed.* Drinking and Intoxication. Glencoe: Free Press, 1959. Pp. 455.
- MERCER, J. D. They Walk in Shadow. New York: Comet Press, 1959. Pp. 573.
- MEYER, H. J., & BORGATTA, E. F. An Experiment in Mental Patient Rehabilitation. New York: Russell Sage Foundation, 1959. Pp. 114.
- MILLER, D. R., & SWANSON, G. E. The Changing American Parent. New York: Wiley, 1958. Pp. 302.
- MORTENSEN, D. G., & SCHMULLER, A. M. Guidance in Today's Schools. New York: Wiley, 1959. Pp. 436.
- MOUSTAKAS, C. E. The Alive and Growing Teacher. New York: Philosophical Library, 1959. Pp. 157.

- MURHERJEE, B. N. *A Laboratory Guide in Psychology*. Patna: Shri Himalaya Press, 1956. Pp. 103.
- MYERS, J. K., & ROBERTS, B. H. *Family and Class Dynamics in Mental Illness*. New York: Wiley, 1959. Pp. 295.
- NEMECEK, O. *Virginity: Pre-nuptial Rites and Rituals*. New York: Philosophical Library, 1958. Pp. 129.
- O'BRIEN, B. *Operators and Things: The Inner Life of a Schizophrenic*. Cambridge: Arlington Books, 1959. Pp. 166.
- OLSON, W. C. *Child Development*. Boston: Heath, 1959. Pp. 497.
- OSTROVSKY, E. S. *Father to the Child*. New York: Putnam, 1959. Pp. 173.
- PARNELL, R. W. *Behaviour and Physique*. London: Edward Arnold, 1958. Pp. 134.
- PATTERSON, C. H. *Counseling and Psychotherapy: Theory and Practice*. New York: Harper, 1959. Pp. 322.
- PRESSEY, S. L., ROBINSON, F. P., & HORROCKS, J. E. *Psychology in Education*. New York: Harper, 1959. Pp. 658.
- RASEY, M. I., *Ed.* *The Nature of Being Human*. Detroit: Wayne State Univ. Press, 1959. Pp. 115.
- REDL, F., & WATTENBERG, W. W. *Mental Hygiene in Teaching*. New York: Harcourt, Brace, 1959. Pp. 562.
- REIK, T. *Of Love and Lust*. New York: Grove, 1959. Pp. 623.
- REISS, S. *Language and Psychology*. New York: Philosophical Library, 1959. Pp. 299.
- RINKEL, M., *Ed.* *Chemical Concepts of Psychosis*. New York: McDowell, Obolensky, 1958. Pp. 485.
- ROBERTSON, J. *Young Children in Hospitals*. New York: Basic Books, 1959. Pp. 136.
- ROTHNEY, J. W. M., DANIELSON, P. J., & HEIMANN, R. A. *Measurement for Guidance*. New York: Harper, 1959. Pp. 378.
- SARASON, S. B. *Psychological Problems in Mental Deficiency*. New York: Harper, 1959. Pp. 678.
- SCHACHTER, S. *The Psychology of Affiliation*. Stanford: Stanford Univ. Press, 1959. Pp. 141.
- SCHAFFNER, B., *Ed.* *Group Processes*. New York: Josiah Macy, Jr. Foundation, 1959. Pp. 266.
- SCOTT, W. A., & WITHEY, S. B. *The United States and the United Nations: The Public View, 1945-1955*. New York: Manhattan, 1958. Pp. 314.
- SEWARD, G. S., & SEWARD, J. P., *Eds.* *Current Psychological Issues*. New York: Henry Holt, 1958. Pp. 360.
- SKINNER, B. F. *Cumulative Record*. New York: Appleton-Century-Crofts, Inc., 1959. Pp. 430.
- STANDAL, S. W., CORSINI, R. J., *Eds.* *Critical Incidents in Psychotherapy*. Englewood Cliffs, N. J. Prentice-Hall, 1959. Pp. 396.
- STOGDILL, R. M. *Individual Behavior and Group Achievement*. New York: Oxford Univ. Press, 1959. Pp. 352.
- STONE, A. A., & ONQUE, G. C. *Longitudinal Studies of Child Personality*. Cambridge: Harvard Univ. Press, 1959. Pp. 314.
- STREAM, L. P. *The Birth of Normal Babies*. New York: Twayne, 1958. Pp. 194.
- SUTHERLAND, J. D., *Ed.* *Psychoanalysis and Contemporary Thought*. New York: Grove, 1959. Pp. 149.
- TASHMAN, H. F. *The Marriage Bed*. New York: University Publishers, 1959. Pp. 303.
- TASHMAN, H. F. *Today's Neurotic Family*. New York: University Publishers, 1959. Pp. 214.
- TAUBER, E. S., & GREEN, M. R. *Prelogical Experience*. New York: Basic Books, 1959. Pp. 196.
- TERMAN, L. M., *Ed.* *Genetic Studies of Genius. The Gifted Group at Mid-Life*. Stanford: Stanford Univ. Press, 1959. Pp. 187.
- THOMAS, W. I., & ZNANIECKI, F. *The Polish Peasant in Europe and America*. New York: Dover, 1958. Pp. 1114. Vol. 1.

- THOMAS, W. I., & ZNANIECKI, F. The Polish Peasant in Europe and America. New York: Dover, 1958. Pp. 1157-2250. Vol. 11.
- THOMPSON, E. T., & HUGHES, E. C., *Eds.* Race: Individual and Collective Behavior. Glencoe: Free Press, 1958. Pp. 619.
- THURSTONE, L. L. The Measurement of Values. Chicago: Univ. Chicago Press, 1959. Pp. 322.
- TRAVERS, R. M. W. An Introduction to Educational Research. New York: Macmillan, 1958. Pp. 466.
- VOGT, E. Z., & HYMAN, R. Water Witching U.S.A. Chicago: Univ. Chicago Press, 1959. Pp. 248.
- WALSHE, F. Diseases of the Nervous System. Baltimore: Williams & Wilkins, 1958. Pp. 373.
- WATSON, R. I. Psychology of the Child. New York: Wiley, 1959. Pp. 662.
- WILMER, H. A. Social Psychiatry in Action. Springfield: Thomas, 1958. Pp. 373.
- WOLF, A. V. Thirst. Springfield: Thomas, 1958. Pp. 536.

